# BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

## IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.**

10AT

**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07H 21/00, C07K 14/00 — A1

(11) International Publication Number: **WO 97/33899**

(43) International Publication Date: 18 September 1997 (18.09.97)

(21) International Application Number: PCT/US96/03773

(22) International Filing Date: 14 March 1996 (14.03.96)

(71) Applicant *(for all designated States except US)*: HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850-3338 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US).

(74) Agent: BENSON, Robert, H.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).

(81) Designated States: AM, AU, BG, BR, BY, CA, CN, CZ, EE, FI, GE, HU, JP, KG, KP, KR, KZ, LT, LV, MD, MN, MX, NO, NZ, PL, RO, RU, SG, SI, SK, TJ, TM, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

**Published**
*With international search report.*

(54) Title: APOPTOSIS INDUCING MOLECULE I

(57) Abstract

The invention relates to AIM-I polypeptides, polynucleotides encoding the polypeptides, methods for producing the polypeptides, in particular by expressing the polynucleotides, and agonists and antagonists of the polypeptides. The invention further relates to methods for utilizing such polynucleotides, polypeptides, agonists and antagonists for applications, which relate, in part, to research, diagnostic and clinical arts.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | |
|---|---|---|---|---|---|
| **AM** | Armenia | **GB** | United Kingdom | **MW** | Malawi |
| **AT** | Austria | **GE** | Georgia | **MX** | Mexico |
| **AU** | Australia | **GN** | Guinea | **NE** | Niger |
| **BB** | Barbados | **GR** | Greece | **NL** | Netherlands |
| **BE** | Belgium | **HU** | Hungary | **NO** | Norway |
| **BF** | Burkina Faso | **IE** | Ireland | **NZ** | New Zealand |
| **BG** | Bulgaria | **IT** | Italy | **PL** | Poland |
| **BJ** | Benin | **JP** | Japan | **PT** | Portugal |
| **BR** | Brazil | **KE** | Kenya | **RO** | Romania |
| **BY** | Belarus | **KG** | Kyrgystan | **RU** | Russian Federation |
| **CA** | Canada | **KP** | Democratic People's Republic of Korea | **SD** | Sudan |
| **CF** | Central African Republic | | | **SE** | Sweden |
| **CG** | Congo | **KR** | Republic of Korea | **SG** | Singapore |
| **CH** | Switzerland | **KZ** | Kazakhstan | **SI** | Slovenia |
| **CI** | Côte d'Ivoire | **LI** | Liechtenstein | **SK** | Slovakia |
| **CM** | Cameroon | **LK** | Sri Lanka | **SN** | Senegal |
| **CN** | China | **LR** | Liberia | **SZ** | Swaziland |
| **CS** | Czechoslovakia | **LT** | Lithuania | **TD** | Chad |
| **CZ** | Czech Republic | **LU** | Luxembourg | **TG** | Togo |
| **DE** | Germany | **LV** | Latvia | **TJ** | Tajikistan |
| **DK** | Denmark | **MC** | Monaco | **TT** | Trinidad and Tobago |
| **EE** | Estonia | **MD** | Republic of Moldova | **UA** | Ukraine |
| **ES** | Spain | **MG** | Madagascar | **UG** | Uganda |
| **FI** | Finland | **ML** | Mali | **US** | United States of America |
| **FR** | France | **MN** | Mongolia | **UZ** | Uzbekistan |
| **GA** | Gabon | **MR** | Mauritania | **VN** | Viet Nam |

# Apoptosis Inducing Molecule I

This invention relates, in part, to newly identified polynucleotides and polypeptides; variants and derivatives of the polynucleotides and polypeptides; processes for making the polynucleotides and the polypeptides, and their variants and derivatives; agonists and antagonists of the polypeptides; and uses of the polynucleotides, polypeptides, variants, derivatives, agonists and antagonists. In particular, in these and in other regards, the invention relates to polynucleotides and polypeptides of human Apoptosis Inducing Molecule I (AIM-I).

## BACKGROUND OF THE INVENTION

Human tumor necrosis factors $\alpha$ (TNF-$\alpha$) and $\beta$ (TNF-$\beta$ or lymphotoxin) are related members of a broad class of polypeptide mediators, which includes the interferons, interleukins and growth factors, collectively called cytokines (Beutler, B. and Cerami, A., *Annu. Rev. Immunol.*, 7:625-655, 1989).

Tumor necrosis factor (TNF-$\alpha$ and TNF-$\beta$) was originally discovered as a result of its anti-tumor activity, however, now it is recognized as a pleiotropic cytokine capable of numerous biological activities including apoptosis of some transformed cell lines, mediation of cell activation and proliferation and also as playing important roles in immune regulation and inflammation.

To date, there are nine known members of the TNF-ligand superfamily, TNF-$\alpha$, TNF-$\beta$ (lymphotoxin-$\alpha$), LT-$\beta$, OX40L, Fas ligand, CD30L, CD27L, CD40L and 4-1BBL. The ligands of the TNF ligand superfamily are acidic, TNF-like molecules with approximately 20% sequence homology in the extracellular domains (range, 12%-36%) and exist mainly as membrane-bound forms with the biologically active form being a trimeric/multimeric complex. Soluble forms of the TNF ligand superfamily have only been identified so far for TNF, LT$\alpha$, and Fas ligand (for a general review, see Gruss, H. and Dower, S.K., *Blood, 85 (12)*:3378-3404, 1995), which is hereby incorporated by reference in its entirety.

These proteins are involved in regulation of cell proliferation, activation, and differentiation, including control of cell survival or death by apoptosis or cytotoxicity (Armitage, R.J., *Curr. Opin. Immunol.*, 6:407, 1994 and Smith, C.A., *Cell, 75*:959, 1994).

Mammalian development is dependent on both the proliferation and differentiation of cells as well as programmed cell death which occurs through apoptosis (Walker *et al*, *Methods Achiev. Exp. Pathol. 13*:18, 1988). Apoptosis plays a critical role in the destruction of immune thymocytes that recognize self antigens. Failure of this normal elimination process may play a role in autoimmune diseases (Gammon *et al.*, *Immunology Today*, *12*:193, 1991).

Itoh *et al.*, *Cell*, 66:233, 1991, described a cell surface antigen, Fas/CD23 that mediates apoptosis and is involved in clonal deletion of T-cells. Fas is expressed in activated T-cells, B-cells, neutrophils and in thymus, liver, heart and lung and ovary in adult mice (Watanabe-Fukunaga *et al.*, *J. Immunol., 148*:1274, 1992) in addition to activated T-cells, B-cells, neutrophils. In experiments where a monoclonal antibody is cross-linked to Fas, apoptosis is induced (Yonehara *et al.*, *J. Exp. Med, 169*:1747, 1989; Trauth *et al.*, *Science*, *245*:301, 1989). In addition, there is an example where

binding of a monoclonal antibody to Fas ligand is stimulatory to T-cells under certain conditions (Alderson *et al.*, *J. Exp. Med* 178:2231, 1993).

Fas antigen is a cell surface protein of relative MW of 45 Kd. Both human and murine genes for Fas have been cloned by Watanabe-Fukunaga *et al.*, (*J. Immunolo.* 148:1274, 1992) and Itoh *et al.* (*Cell, 66*:233, 1991). The proteins encoded by these genes are both transmembrane proteins with structural homology to the nerve growth factor/tumor necrosis factor receptor superfamily, which includes two TNF receptors, the low affinity nerve growth factor receptor and CD40, CD27, CD30, and OX40.

Recently the Fas ligand has been described (Suda *et al.*, *Cell, 75*:1169, 1993). The amino acid sequence indicates that Fas ligand is a type 11 transmembrane protein belonging to the TNF family. Fas ligand is expressed in splenocytes and thymocytes, consistent with T-cell mediated cytotoxicity. The purified Fas ligand has a MW of 40 Kd.

Recently it has been demonstrated that Fas/Fas ligand interactions are required for apoptosis following the activation of T-cells (Ju *et al. Nature, 373:444,* 1995; Brunner *et al.*, *Nature, 373*:441, 1995). Activation of T-cells induces expression of both proteins on the cell surface. Subsequent interaction between the ligand and receptor results in apoptosis of the cells. This supports the possible regulatory role for apoptosis induced by Fas/Fas ligand interaction during normal immune responses.

Clearly, there is a need for factors that regulate activation, and differentiation of normal and abnormal cells. There is a need, therefore, for identification and characterization of such factors that modulate activation and differentiation of cells, both normally and in disease states. In particular, there is a need to isolate and characterize additional molecules that mediate apoptosis, and treat autoimmune disease, graft versus host disease and lymphadenopathy, and, among other things, can play a role in preventing, ameliorating or correcting dysfunctions or diseases.

## SUMMARY OF THE INVENTION

The polypeptide of the present invention has been identified as a novel member of the TNF ligand super-family based on structural and biological similarities.

Toward these ends, and others, it is an object of the present invention to provide polypeptides, *inter alia*, that have been identified as novel Apoptosis Inducing Molecule I (AIM-I) polypeptides by homology between the amino acid sequence set out in Figure 1 and known amino acid sequences of other proteins such as human Fas ligand (Suda *et al.*, *Cell*, *75*:1169, 1993).

It is a further object of the invention, moreover, to provide polynucleotides that encode AIM-I, particularly polynucleotides that encode the polypeptide herein designated AIM-I. In a particularly preferred embodiment of this aspect of the invention the polynucleotide comprises the region encoding human AIM-I in the sequence set out in Figure 1. In accordance with this aspect of the present invention there is provided an isolated nucleic acid molecule encoding a mature polypeptide expressed by the human cDNA contained in ATCC Deposit No. 97448.

In accordance with this aspect of the invention there are provided isolated nucleic acid molecules encoding human AIM-I, including mRNAs, cDNAs, genomic DNAs and, in further embodiments of this aspect of the invention, biologically, diagnostically, clinically or therapeutically useful variants, analogs, derivatives and/or fragments thereof, including fragments of the variants, analogs and derivatives. Among the particularly preferred embodiments of this aspect of the invention are naturally occurring allelic variants of human AIM-I

It also is an object of the invention to provide AIM-I polypeptides, particularly human AIM-I polypeptides, which may be employed to treat lymphadenopathy,

autoimmune disease, graft versus host disease; which may be used to stimulate peripheral tolerance, destroy pathologic transformed cell lines, mediate cell activation and proliferation; and are functionally linked as primary mediators of immune regulation and inflammatory response.

This aspect of the invention provides novel polypeptides of human origin referred to herein as AIM-I as well as biologically, diagnostically or therapeutically useful fragments, variants and derivatives thereof, variants and derivatives of the fragments, and analogs of all of the foregoing. Among the particularly preferred embodiments of this aspect of the invention are variants of human AIM-I encoded by naturally occurring alleles of the human AIM-I gene.

It is another object of the invention to provide a process for producing the aforementioned polypeptides, polypeptide fragments, variants and derivatives, fragments of the variants and derivatives, and analogs of the foregoing. In a preferred embodiment of this aspect of the invention there are provided methods for producing the aforementioned AIM-I polypeptides comprising culturing host cells having expressibly incorporated therein an exogenously-derived human AIM-I-encoding polynucleotide under conditions for expression of human AIM-I in the host and then recovering the expressed polypeptide.

In accordance with another object the invention there are provided products, compositions, processes and methods that utilize the aforementioned polypeptides and polynucleotides for research, biological, clinical and therapeutic purposes, *inter alia*.

In accordance with certain preferred embodiments of this aspect of the invention, there are provided products, compositions and methods, *inter alia*, for, among other things: assessing AIM-I expression in cells by determining AIM-I polypeptides or AIM-I-encoding mRNA; treating disease or disorder caused by under-expression of the AIM-I *in vitro*, *ex vivo* or *in vivo* by exposing cells to AIM-I polypeptides or

polynucleotides as disclosed herein; assaying genetic variation and aberrations, such as defects, in AIM-I genes; and administering a AIM-I polypeptide or polynucleotide to an organism to augment AIM-I function or remediate AIM-I dysfunction.

In accordance with certain preferred embodiments of these and other aspects of the invention there are provided probes that hybridize to human AIM-I sequences.

Certain additional preferred aspects related to the above aspects of the invention provides antibodies against AIM-I polypeptides. In particularly preferred embodiments in this regard, the antibodies are highly selective for human AIM-I, and may be employed, *inter alia*, to treat autoimmune diseases.

In accordance with another aspect of the present invention, there are provided AIM-I agonists. Among preferred agonists are molecules that mimic AIM-I, that bind to AIM-I-binding molecules or receptor molecules, and that elicit or augment AIM-I-induced responses. Also among preferred agonists are molecules that interact with AIM-I or AIM-I polypeptides, or with other modulators of AIM-I activities, and thereby potentiate or augment an effect of AIM-I or more than one effect of AIM-I.

In accordance with yet another aspect of the present invention, there are provided AIM-I antagonists. Among preferred antagonists are those which mimic AIM-I so as to bind to AIM-I receptor or binding molecules but not elicit an AIM-I-induced response or more than one AIM-I-induced response. Also among preferred antagonists are molecules that bind to or interact with AIM-I so as to inhibit an effect of AIM-I or more than one effect of AIM-I or which prevent expression of AIM-I. The antagonists may be employed to prevent septic shock, inflammation, cerebral malaria, activation of the HIV virus, graft-host rejection, bone resorption, rheumatoid arthritis and cachexia.

In a further aspect of the invention there are provided compositions comprising an AIM-I polynucleotide or an AIM-I polypeptide for administration to cells *in vitro*,

to cells *ex vivo* and to cells *in vivo*, or to a multicellular organism. In certain particularly preferred embodiments of this aspect of the invention, the compositions comprise an AIM-I polynucleotide for expression of an AIM-I polypeptide in a host organism for treatment of disease. Particularly preferred in this regard is expression in a human patient for treatment of a dysfunction associated with aberrant endogenous activity of AIM-I.

Other objects, features, advantages and aspects of the present invention will become apparent to those of skill from the following description. It should be understood, however, that the following description and the specific examples, while indicating preferred embodiments of the invention, are given by way of illustration only. Various changes and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the following description and from reading the other parts of the present disclosure.

## BRIEF DESCRIPTION OF THE DRAWINGS

The following drawings depict certain embodiments of the invention. They are illustrative only and do not limit the invention otherwise disclosed herein.

Figure 1 shows the nucleotide and deduced amino acid sequence of human AIM-1.

Figure 2 shows the regions of similarity and identity between amino acid sequence of AIM-I of the present invention and human Fas ligand polypeptide (SEQ ID NO:3).

Figure 3 shows regions of similarity and identity between amino acid sequences of AIM-I of the present invention and human Fas ligand polypeptide, TNF$\alpha$ and TNF$\beta$.

Figure 4 shows structural and functional features of AIM-I deduced by the indicated techniques, as a function of amino acid sequence.

## DETAILED DESCRIPTION OF THE INVENTION

The following illustrative elucidations are provided to facilitate understanding of certain terms used frequently herein, particularly in the examples. The elucidations are provided as a convenience and are not limitative of the invention.

The term "digestion of DNA" refers to catalytic cleavage of DNA with a restriction enzyme that acts only upon certain sequences in the DNA. The various restriction enzymes referred to herein are all commercially available and their reaction conditions, cofactors and other requirements for use are known and routine to the skilled artisan.

For analytical purposes, typically, 1 μg of plasmid or DNA fragment is digested with about 2 units of enzyme in about 20 μl of reaction buffer. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50 μg of DNA are digested with 20 to 250 units of enzyme in proportionately larger volumes.

Appropriate buffers and substrate amounts for particular restriction enzymes are described in standard laboratory manuals, such as those referenced below, and they are specified by commercial suppliers.

Incubation times of about 1 hour at 37°C are ordinarily used, but conditions may vary in accordance with standard procedures, the supplier's instructions and the particulars of the reaction. After digestion, reactions may be analyzed, and fragments may be purified by electrophoresis through an agarose or polyacrylamide gel, using well known methods that are routine for those skilled in the art.

The term "genetic element" generally means a polynucleotide comprising a region that encodes a polypeptide or a region that regulates transcription or translation or other processes important to expression of the polypeptide in a host cell, or a polynucleotide comprising both a region that encodes a polypeptide and a region operably linked thereto that regulates expression.

Genetic elements may be comprised within a vector that replicates as an episomal element; that is, as a molecule physically independent of the host cell genome. They may be comprised within mini-chromosomes, such as those that arise during amplification of transfected DNA by methotrexate selection in eukaryotic cells. Genetic elements also may be comprised within a host cell genome; not in their natural state but, rather, following manipulation such as isolation, cloning and introduction into a host cell in the form of purified DNA or in a vector, among others.

The term "isolated" means altered "by the hand of man" from its natural state; *i.e.*, that, if it occurs in nature, it has been changed or removed from its original environment, or both.

For example, a naturally occurring polynucleotide or a polypeptide naturally present in a living animal in its natural state is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. For example, with respect to polynucleotides, the term isolated means that it is separated from the chromosome and cell in which it naturally occurs

As part of or following isolation, such polynucleotides can be joined to other polynucleotides, such as DNAs, for mutagenesis, to form fusion proteins, and for propagation or expression in a host, for instance. The isolated polynucleotides, alone or joined to other polynucleotides such as vectors, can be introduced into host cells, in culture or in whole organisms. Introduced into host cells in culture or in whole

organisms, such DNAs still would be isolated, as the term is used herein, because they would not be in their naturally occurring form or environment. Similarly, the polynucleotides and polypeptides may occur in a composition, such as a media formulations, solutions for introduction of polynucleotides or polypeptides, for example, into cells, compositions or solutions for chemical or enzymatic reactions, for instance, which are not naturally occurring compositions, and, therein remain isolated polynucleotides or polypeptides within the meaning of that term as it is employed herein.

The term "ligation" refers to the process of forming phosphodiester bonds between two or more polynucleotides, which most often are double stranded DNAs. Techniques for ligation are well known to the art and protocols for ligation are described in standard laboratory manuals and references, such as, for instance, Sambrook *et al.*, Molecular Cloning. A Laboratory Manual. 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989) and Maniatis *et al.*, pg. 146, as cited below.

The tern "oligonucleotide(s)" refers to relatively short polynucleotides. Often the term refers to single-stranded deoxyribonucleotides, but it can refer as well to single- or double-stranded ribonucleotides. RNA:DNA hybrids and double-stranded DNAs, among others.

Oligonucleotides, such as single-stranded DNA probe oligonucleotides, often are synthesized by chemical methods, such as those implemented on automated oligonucleotide synthesizers. However, oligonucleotides can be made by a variety of other methods, including *in vitro* recombinant DNA-mediated techniques and by expression of DNAs in cells and organisms.

Initially, chemically synthesized DNAs typically are obtained without a 5' phosphate. The 5' ends of such oligonucleotides are not substrates for phosphodiester bond formation by ligation reactions that employ DNA ligases typically used to form

recombinant DNA molecules. Where ligation of such oligonucleotides is desired, a phosphate can be added by standard techniques, such as those that employ a kinase and ATP.

The 3' end of a chemically synthesized oligonucleotide generally has a free hydroxyl group and, in the presence of a ligase, such as T4 DNA ligase, readily will form a phosphodiester bond with a 5' phosphate of another polynucleotide, such as another oligonucleotide. As is well known, this reaction can be prevented selectively, where desired, by removing the 5' phosphates of the other polynucleotide(s) prior to ligation.

The term "plasmids" generally are designated herein by a lower case p preceded and/or followed by capital letters and/or numbers, in accordance with standard naming conventions that are familiar to those of skill in the art.

Starting plasmids disclosed herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids by routine application of well known, published procedures. Many plasmids and other cloning and expression vectors that can be used in accordance with the present invention are well known and readily available to those of skill in the art. Moreover, those of skill readily may construct any number of other plasmids suitable for use in the invention. The properties, construction and use of such plasmids, as well as other vectors, in the present invention will be readily apparent to those of skill from the present disclosure.

The term "polynucleotide(s)" generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. Thus, for instance, polynucleotides as used herein refers to, among others, single-and double-stranded DNA, DNA that is a mixture of single-and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and

double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions.

In addition, polynucleotide as used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide.

As used herein, the term polynucleotide includes DNAs or RNAs as described above that contain one or more modified bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "polynucleotides" as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein.

It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term polynucleotide as it is employed herein embraces such chemically, enzymatically or metabolically modified forms of polynucleotides, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including simple and complex cells, *inter alia*.

The term "polypeptides" as used herein, includes all polypeptides as described below. The basic structure of polypeptides is well known and has been described in innumerable textbooks and other publications in the art. In this context, the term is used herein to refer to any peptide or protein comprising two or more amino acids joined to each other in a linear chain by peptide bonds. As used herein, the term refers to both short chains, which also commonly are referred to in the art as peptides, oligopeptides

and oligomers, for example, and to longer chains, which generally are referred to in the art as proteins, of which there are many types.

It will be appreciated that polypeptides often contain amino acids other than the 20 amino acids commonly referred to as the 20 naturally occurring amino acids, and that many amino acids, including the terminal amino acids, may be modified in a given polypeptide, either by natural processes, such as processing and other post-translational modifications, but also by chemical modification techniques which are well known to the art. Even the common modifications that occur naturally in polypeptides are too numerous to list exhaustively here, but they are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature, and they are well known to those of skill in the art. Among the known modifications which may be present in polypeptides of the present are, to name an illustrative few, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

Such modifications are well known to those of skill and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, for instance, are described in most basic texts, such as, for instance Proteins-Structure And Molecular Properties, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993). Many detailed reviews are available on this subject, such as, for example, those provided by Wold, F.,

Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in Posttranslational Covalent Modification Of Proteins, B. C. Johnson, Ed., Academic Press, New York (1983); Seifter *et al.*, Analysis for protein modifications and nonprotein cofactors, *Meth. Enzymol.*, *182*:626-646 (1990) and Rattan *et al.*, Protein Synthesis: Posttranslational Modifications and Aging, *Ann. N.Y. Acad. Sci.*, *663*:48-62 (1992).

It will be appreciated, as is well known and as noted above, that polypeptides are not always entirely linear. For instance, polypeptides may be branched as a result of ubiquitination, and they may be circular, with or without branching, generally as a result of posttranslation events, including natural processing event and events brought about by human manipulation which do not occur naturally. Circular, branched and branched circular polypeptides may be synthesized by non-translation natural process and by entirely synthetic methods, as well.

Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. In fact, blockage of the amino or carboxyl group in a polypeptide, or both, by a covalent modification, is common in naturally occurring and synthetic polypeptides and such modifications may be present in polypeptides of the present invention, as well. For instance, the amino terminal residue of polypeptides made in *E. coli*, prior to proteolytic processing, almost invariably will be N-formylmethionine.

The modifications that occur in a polypeptide often will be a function of how it is made. For polypeptides made by expressing a cloned gene in a host, for instance, the nature and extent of the modifications in large part will be determined by the host cell posttranslational modification capacity and the modification signals present in the polypeptide amino acid sequence. For instance, as is well known, glycosylation often does not occur in bacterial hosts such as *E. coli*. Accordingly, when glycosylation is desired, a polypeptide should be expressed in a glycosylating host, generally a

eukaryotic cell. Insect cell often carry out the same posttranslational glycosylations as mammalian cells and, for this reason, insect cell expression systems have been developed to express efficiently mammalian proteins having native patterns of glycosylation, *inter alia*. Similar considerations apply to other modifications.

It will be appreciated that the same type of modification may be present in the same or varying degree at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications.

In general, as used herein, the term polypeptide encompasses all such modifications, particularly those that are present in polypeptides synthesized by expressing a polynucleotide in a host cell.

The term "variant(s)"of polynucleotides or polypeptides, as the term is used herein, are polynucleotides or polypeptides that differ from a reference polynucleotide or polypeptide, respectively. Variants in this sense are described below and elsewhere in the present disclosure in greater detail.

A polynucleotide variant can be, for example, a polynucleotide that differs in nucleotide sequence from another, reference polynucleotide. Generally, differences are limited so that the nucleotide sequences of the reference and the variant are closely similar overall and, in many regions, identical.

As noted below, changes in the nucleotide sequence of the variant may be silent. That is, they may not alter the amino acids encoded by the polynucleotide. Where alterations are limited to silent changes of this type a variant will encode a polypeptide with the same amino acid sequence as the reference. Also as noted below, changes in the nucleotide sequence of the variant may alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Such nucleotide changes may

result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below.

A polynucleotide variant can be, for example, a polypeptide that differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference and the variant are closely similar overall and, in many region, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions, fusions and truncations, which may be present in any combination.

The term "receptor molecule" as used herein, refers to molecules which bind or interact specifically with AIM-I polypeptides of the present invention, including not only classic receptors, which are preferred, but also other molecules that specifically bind to or interact with polypeptides of the invention (which also may be referred to as "binding molecules" and "interaction molecules," respectively and as "AIM-I binding molecules" and "AIM-I interaction molecules.") Binding between polypeptides of the invention and such molecules, including receptor or binding or interaction molecules may be exclusive to polypeptides of the invention, which is very highly preferred, or it may be highly specific for polypeptides of the invention, which is highly preferred, or it may be highly specific to a group of proteins that includes polypeptides of the invention, which is preferred, or it may be specific to several groups of proteins at least one of which includes polypeptides of the invention.

Receptors also may be non-naturally occurring, such as antibodies and antibody-derived reagents that bind specifically to polypeptides of the invention.

The present invention relates to novel AIM-I polypeptides and polynucleotides, among other things, as described in greater detail below. In particular, the invention relates to polypeptides and polynucleotides of a novel human AIM-I, which is related by amino acid sequence homology to known human AIM-I. The invention relates

especially to AIM-I having the nucleotide and amino acid sequences set out in Figure 1, and to the AIM-I nucleotide and amino acid sequences of the cDNA in ATCC Deposit No. 97448, which is herein referred to as "the deposited clone" or as the "cDNA of the deposited clone." It will be appreciated that the nucleotide and amino acid sequences set out in Figure 1 were obtained by sequencing the cDNA of the deposited clone. Hence, the sequence of the deposited clone is controlling as to any discrepancies between the two and any reference to the sequences of Figure 1 include reference to the sequence of the human cDNA of the deposited clone.

In accordance with one aspect of the present invention, there are provided isolated polynucleotides which encode the AIM-I polypeptide having the deduced amino acid sequence of Figure 1 or the AIM-I polypeptide encoded by the cDNA in the deposited clone.

Using the information provided herein, such as the polynucleotide sequence set out in Figure 1, a polynucleotide of the present invention encoding human AIM-I polypeptide may be obtained using standard cloning and screening procedures, such as those for cloning cDNAs using human tissue as starting material. Illustrative of the invention, the polynucleotide set out in Figure 1 was discovered in a cDNA library derived from cells of a human pancreas tumor.

Human AIM-I of the invention is structurally related to other proteins of the TNF family, as shown by the results of sequencing the human cDNA encoding human AIM-I in the deposited clone. The cDNA sequence thus obtained is set out in Figure 1. It contains an open reading frame encoding a protein of about 281 amino acid residues with a deduced molecular weight of about 31 kDa. The protein exhibits greatest homology to known human AIM-I, among known proteins. The AIM-I of Figure 1 has about 22.892 % identity and about 48.594 % similarity with the amino acid sequence of human Fas ligand.

Polynucleotides of the present invention may be in the form of RNA, such as mRNA, or in the form cDNA and genomic DNA obtained by cloning or produced by chemical synthetic techniques or by a combination thereof. The DNA may be double-stranded or single-stranded. Single-stranded DNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

The coding sequence which encodes the polypeptide may be identical to the coding sequence of the polynucleotide shown in Figure 1. It also may be a polynucleotide with a different sequence, which, as a result of the redundancy (degeneracy) of the genetic code, encodes the polypeptide of the DNA of Figure 1.

Polynucleotides of the present invention which encode the polypeptide of Figure 1 may include, but are not limited to the coding sequence for the mature polypeptide, by itself; the coding sequence for the mature polypeptide and additional coding sequences, such as those encoding a leader or secretory sequence, such as a pre-, or pro- or prepro- protein sequence; the coding sequence of the mature polypeptide, with or without the aforementioned additional coding sequences, together with additional, non-coding sequences, including for example, but not limited to introns and non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences that play a role in transcription, mRNA processing - including splicing and polyadenylation signals, for example - ribosome binding and stability of mRNA; additional coding sequence which codes for additional amino acids, such as those which provide additional functionalities. Thus, for instance, the polypeptide may be fused to a marker sequence, such as a peptide, which facilitates purification of the fused polypeptide. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, such as the tag provided in the pQE vector (Qiagen, Inc., Chatsworth, CA), among others, many of which are commercially available. As described in Gentz *et al.*, *Proc. Natl. Acad. Sci., USA, 86*:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein.

The HA tag corresponds to an epitope derived of influenza hemagglutinin protein, which has been described by Wilson *et al.*, *Cell. 37*:767 (1984), for instance.

In accordance with the foregoing, the term "polynucleotide encoding a polypeptide" as used herein encompasses polynucleotides which include a sequence encoding a polypeptide of the present invention, particularly the human AIM-I having the amino acid sequence set out in Figure 1. The term encompasses polynucleotides that include a single continuous region or discontinuous regions encoding the polypeptide (for example, interrupted by introns) together with additional regions, that also may contain coding and/or non-coding sequences.

The present invention further relates to variants of the herein above described polynucleotides which encode for fragments, analogs and derivatives of the polypeptide having the deduced amino acid sequence of Figure 1. A variant of the polynucleotide may be a naturally occurring variant such as a naturally occurring allelic variant, or it may be a variant that is not known to occur naturally. Such non-naturally occurring variants of the polynucleotide may be made by mutagenesis techniques, including those applied to polynucleotides, cells or organisms.

Among variants in this regard are variants that differ from the aforementioned polynucleotides by nucleotide substitutions, deletions or additions. The substitutions, deletions or additions may involve one or more nucleotides. The variants may be altered in coding or non-coding regions or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or additions.

Among the particularly preferred embodiments of the invention in this regard are polynucleotides encoding polypeptides having the amino acid sequence of AIM-I set out in Figure 1; variants, analogs, derivatives and fragments thereof, and fragments of the variants, analogs and derivatives.

Further particularly preferred in this regard are polynucleotides encoding AIM-I variants, analogs, derivatives and fragments, and variants, analogs and derivatives of the fragments, which have the amino acid sequence of the AIM-I polypeptide of Figure 1 in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, deleted or added, in any combination. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the AIM-I. Also especially preferred in this regard are conservative substitutions. Most highly preferred are polynucleotides encoding polypeptides having the amino acid sequence of Figure 1 without substitutions.

Further preferred embodiments of the invention are polynucleotides that are at least 70% identical to a polynucleotide encoding the AIM-I polypeptide having the amino acid sequence set out in Figure 1, and polynucleotides which are complementary to such polynucleotides. Alternatively, most highly preferred are polynucleotides that comprise a region that is at least 80% identical to a polynucleotide encoding the AIM-I polypeptide of the cDNA of the deposited clone and polynucleotides complementary thereto. In this regard, polynucleotides at least 90% identical to the same are particularly preferred, and among these particularly preferred polynucleotides, those with at least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred among those with at least 95%, and among these those with at least 98% and at least 99% are particularly highly preferred, with at least 99% being the more preferred.

Particularly preferred embodiments in this respect, moreover, are polynucleotides which encode polypeptides which retain substantially the same biological function or activity as the mature polypeptide encoded by the cDNA of Figure 1.

The present invention further relates to polynucleotides that hybridize to the herein above-described sequences. In this regard, the present invention especially relates to polynucleotides which hybridize under stringent conditions to the herein above-

described polynucleotides. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences.

As discussed additionally herein regarding polynucleotide assays of the invention, for instance, polynucleotides of the invention as discussed above, may be used as a hybridization probe for cDNA and genomic DNA to isolate full-length cDNAs and genomic clones encoding human AIM-I and to isolate cDNA and genomic clones of other genes that have a high sequence similarity to the human AIM-I gene. Such probes generally will comprise at least 15 bases. Preferably, such probes will have at least 30 bases and may have at least 50 bases. Particularly preferred probes will have at least 30 bases and will have 50 bases or less.

For example, the coding region of the AIM-I gene may be isolated by screening using the known DNA sequence to synthesize an oligonucleotide probe. A labeled oligonucleotide having a sequence complementary to that of a gene of the present invention is then used to screen a library of human cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

The polynucleotides and polypeptides of the present invention may be employed as research reagents and materials for discovery of treatments and diagnostics to human disease, as further discussed herein relating to polynucleotide assays, *inter alia*.

The polynucleotides may encode a polypeptide which is the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, may facilitate protein trafficking, may prolong or shorten protein half-life or may facilitate manipulation of a protein for assay or production, among other things. As

generally is the case in situ, the additional amino acids may be processed away from the mature protein by cellular enzymes.

A precursor protein, having the mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. When prosequences are removed such inactive precursors generally are activated. Some or all of the prosequences may be removed before activation. Generally, such precursors are called proproteins.

In sum, a polynucleotide of the present invention may encode a mature protein, a mature protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences which are not the leader sequences of a preprotein, or a preproprotein, which is a precursor to a proprotein, having a leader sequence and one or more prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

A deposit containing a human AIM-I cDNA has been deposited with the American Type Culture Collection, as noted above. Also as noted above, the human cDNA deposit is referred to herein as "the deposited clone" or as "the cDNA of the deposited clone."

The deposited clone was deposited with the American Type Culture Collection, 12301 Park Lawn Drive, Rockville, Maryland 20852, USA, on February 20, 1996, and assigned ATCC Deposit No. 97448.

The deposited material is a pBluescript SK (-) plasmid (Stratagene, La Jolla, CA) that contains the full length AIM-I cDNA, referred to as "PF261" upon deposit.

The deposit has been made under the terms of the Budapest Treaty on the international recognition of the deposit of micro-organisms for purposes of patent procedure. The strain will be irrevocably and without restriction or condition released to the public upon the issuance of a patent. The deposit is provided merely as convenience to those of skill in the art and is not an admission that a deposit is required for enablement, such as that required under 35 U.S.C. §112. The sequence of the polynucleotides contained in the deposited material, as well as the amino acid sequence of the polypeptide encoded thereby, are controlling in the event of any conflict with any description of sequences herein. A license may be required to make, use or sell the deposited materials, and no such license is hereby granted.

The present invention further relates to a human AIM-I polypeptide which has the deduced amino acid sequence of Figure 1.

The invention also relates to fragments, analogs and derivatives of these polypeptides. The terms "fragment," "derivative" and "analog" when referring to the polypeptide of Figure 1 means a polypeptide which retains essentially the same biological function or activity as such polypeptide. Thus, an analog includes a proprotein which can be activated by cleavage of the proprotein portion to produce an active mature polypeptide.

The polypeptide of the present invention may be a recombinant polypeptide, a natural polypeptide or a synthetic polypeptide. In certain preferred embodiments it is a recombinant polypeptide.

The fragment, derivative or analog of the polypeptide of Figure 1 may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group,

or (iii) one in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature polypeptide, such as a leader or secretory sequence or a sequence which is employed for purification of the mature polypeptide or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

Among the particularly preferred embodiments of the invention in this regard are polypeptides having the amino acid sequence of AIM-I set out in Figure 1, variants, analogs, derivatives and fragments thereof, and variants, analogs and derivatives of the fragments. Alternatively, particularly preferred embodiments of the invention in this regard are polypeptides having the amino acid sequence of the AIM-I of the cDNA in the deposited clone, variants, analogs, derivatives and fragments thereof, and variants, analogs and derivatives of the fragments.

Among preferred variants are those that vary from a reference by conservative amino acid substitutions. Such substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe, Tyr.

Further particularly preferred in this regard are variants, analogs, derivatives and fragments, and variants, analogs and derivatives of the fragments, having the amino acid sequence of the AIM-I polypeptide of Figure 1 in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, deleted or added, in any combination. Especially preferred among these are silent substitutions, additions and

deletions, which do not alter the properties and activities of the AIM-I. Also especially preferred in this regard are conservative substitutions. Most highly preferred are polypeptides having the amino acid sequence of Figure 1 without substitutions.

The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

The polypeptides of the present invention include the polypeptide of SEQ ID NO:2 (in particular the mature polypeptide) as well as polypeptides which have at least 70% similarity (preferably at least 70% identity) to the polypeptide of SEQ ID NO:2 and more preferably at least 90% similarity (more preferably at least 90% identity) to the polypeptide of SEQ ID NO:2 and still more preferably at least 95% similarity (still more preferably at least 95% identity) to the polypeptide of SEQ ID NO:2 and also include portions of such polypeptides with such portion of the polypeptide generally containing at least 30 amino acids and more preferably at least 50 amino acids.

As known in the art "similarity" between two polypeptides is determined by comparing the amino acid sequence and its conserved amino acid substitutes of one polypeptide to the sequence of a second polypeptide.

Fragments or portions of the polypeptides of the present invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, the fragments may be employed as intermediates for producing the full-length polypeptides. Fragments or portions of the polynucleotides of the present invention may be used to synthesize full-length polynucleotides of the present invention.

Also among preferred embodiments of this aspect of the present invention are polypeptides comprising fragments of AIM-I, most particularly fragments of the AIM-I having the amino acid set out in Figure 1, and fragments of variants and derivatives of the AIM-I of Figure 1. In this regard a fragment is a polypeptide having an amino acid

sequence that entirely is the same as part but not all of the amino acid sequence of the aforementioned AIM-I polypeptides and variants or derivatives thereof.

Such fragments may be "free-standing," *i.e.*, not part of or fused to other amino acids or polypeptides, or they may be comprised within a larger polypeptide of which they form a part or region. When comprised within a larger polypeptide, the presently discussed fragments most preferably form a single continuous region. However, several fragments may be comprised within a single larger polypeptide. For instance, certain preferred embodiments relate to a fragment of an AIM-I polypeptide of the present comprised within a precursor polypeptide designed for expression in a host and having heterologous pre and pro-polypeptide regions fused to the amino terminus of the AIM-I fragment and an additional region fused to the carboxyl terminus of the fragment. Therefore, fragments in one aspect of the meaning intended herein, refers to the portion or portions of a fusion polypeptide or fusion protein derived from AIM-I.

As representative examples of polypeptide fragments of the invention, there may be mentioned those which have from about 100 to about 281 amino acids.

In this context, "about" includes the particularly recited range and ranges larger or smaller by several, a few, 5, 4, 3, 2 or 1 amino acid at either extreme or at both extremes. For instance, about 100-281 amino acids in this context means a polypeptide fragment of 100 plus or minus several, a few, 5, 4, 3, 2 or 1 amino acids to 281 plus or minus several a few, 5, 4, 3, 2 or 1 amino acid residues, *i.e.*, ranges as broad as 100 minus several amino acids to 281 plus several amino acids to as narrow as 100 plus several amino acids to 281 minus several amino acids.

Highly preferred in this regard are the recited ranges plus or minus as many as 5 amino acids at either or at both extremes. Particularly highly preferred are the recited ranges plus or minus as many as 3 amino acids at either or at both the recited extremes. Especially particularly highly preferred are ranges plus or minus 1 amino acid at either

or at both extremes or the recited ranges with no additions or deletions. Most highly preferred of all in this regard are fragments from about 100 to about 281 amino acids.

Among especially preferred fragments of the invention are truncation mutants of AIM-I. Truncation mutants include AIM-I polypeptides having the amino acid sequence of Figure 1, or of variants or derivatives thereof, except for deletion of a continuous series of residues (that is, a continuous region, part or portion) that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus or, as in double truncation mutants, deletion of two continuous series of residues, one including the amino terminus and one including the carboxyl terminus. Fragments having the size ranges set out about also are preferred embodiments of truncation fragments, which are especially preferred among fragments generally.

Also preferred in this aspect of the invention are fragments characterized by structural or functional attributes of AIM-I. Preferred embodiments of the invention in this regard include fragments that comprise alpha-helix and alpha-helix forming regions ("alpha-regions"), beta-sheet and beta-sheet-forming regions ("beta-regions"), turn and turn-forming regions ("turn-regions"), coil and coil-forming regions ("coil-regions"), hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions and high antigenic index regions of AIM-I.

Certain preferred regions in these regards are set out in Figure 3, and include, but are not limited to, regions of the aforementioned types identified by analysis of the amino acid sequence set out in Figure 1. As set out in Figure 3, such preferred regions include Garnier-Robson alpha-regions, beta-regions, turn-regions and coil-regions, Chou-Fasman alpha-regions, beta-regions and turn-regions, Kyte-Doolittle hydrophilic regions and hydrophilic regions, Eisenberg alpha and beta amphipathic regions, Karplus-Schulz flexible regions, Emini surface-forming regions and Jameson-Wolf high antigenic index regions.

Among highly preferred fragments in this regard are those that comprise regions of AIM-I that combine several structural features, such as several of the features set out above. In this regard, the regions defined by the residues about 1 to about 281 of Figure 1, which all are characterized by amino acid compositions highly characteristic of turn-regions, hydrophilic regions, flexible-regions, surface-forming regions, and high antigenic index-regions, are especially highly preferred regions. Such regions may be comprised within a larger polypeptide or may be by themselves a preferred fragment of the present invention, as discussed above. It will be appreciated that the term "about" as used in this paragraph has the meaning set out above regarding fragments in general. Further, a highly preferred fragment comprises amino acids 39 through 281 which constitute a soluble portion of the overall AIM-I polypeptide sequence of Figure 1.

Further preferred regions are those that mediate activities of AIM-I. Most highly preferred in this regard are fragments that have a chemical, biological or other activity of AIM-I, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Highly preferred in this regard are fragments that contain regions that are homologs in sequence, or in position, or in both sequence and to active regions of related polypeptides, such as the related polypeptides set out in Figure 2, which include members of the TNF family. Among particularly preferred fragments in these regards are truncation mutants, as discussed above.

It will be appreciated that the invention also relates to, among others, polynucleotides encoding the aforementioned fragments, polynucleotides that hybridize to polynucleotides encoding the fragments, particularly those that hybridize under stringent conditions, and polynucleotides, such as PCR primers, for amplifying polynucleotides that encode the fragments. In these regards, preferred polynucleotides are those that correspondent to the preferred fragments, as discussed above.

The present invention also relates to vectors which include polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques.

Host cells can be genetically engineered to incorporate polynucleotides and express polypeptides of the present invention. For instance, polynucleotides may be introduced into host cells using well known techniques of infection, transduction, transfection, transvection and transformation. The polynucleotides may be introduced alone or with other polynucleotides. Such other polynucleotides may be introduced independently, co-introduced or introduced joined to the polynucleotides of the invention.

Thus, for instance, polynucleotides of the invention may be transfected into host cells with another, separate, polynucleotide encoding a selectable marker, using standard techniques for co-transfection and selection in, for instance, mammalian cells. In this case the polynucleotides generally will be stably incorporated into the host cell genome.

Alternatively, the polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. The vector construct may be introduced into host cells by the aforementioned techniques. Generally, a plasmid vector is introduced as DNA in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. Electroporation also may be used to introduce polynucleotides into a host. If the vector is a virus, it may be packaged *in vitro* or introduced into a packaging cell and the packaged virus may be transduced into cells. A wide variety of techniques suitable for making polynucleotides and for introducing polynucleotides into cells in accordance with this aspect of the invention are well known and routine to those of skill in the art. Such techniques are reviewed at length in Sambrook *et al.* cited above, which is illustrative of the many laboratory manuals that detail these techniques. In accordance with this aspect of the invention the vector may be, for example, a plasmid vector, a single or double-stranded phage vector, a single

or double-stranded RNA or DNA viral vector. Such vectors may be introduced into cells as polynucleotides, preferably DNA, by well known techniques for introducing DNA and RNA into cells. The vectors, in the case of phage and viral vectors also may be and preferably are introduced into cells as packaged or encapsidated virus by well known techniques for infection and transduction. Viral vectors may be replication competent or replication defective. In the latter case viral propagation generally will occur only in complementing host cells.

Preferred among vectors, in certain respects, are those for expression of polynucleotides and polypeptides of the present invention. Generally, such vectors comprise *cis*-acting control regions effective for expression in a host operatively linked to the polynucleotide to be expressed. Appropriate *trans*-acting factors either are supplied by the host, supplied by a complementing vector or supplied by the vector itself upon introduction into the host.

In certain preferred embodiments in this regard, the vectors provide for specific expression. Such specific expression may be inducible expression or expression only in certain types of cells or both inducible and cell-specific. Particularly preferred among inducible vectors are vectors that can be induced for expression by environmental factors that are easy to manipulate, such as temperature and nutrient additives. A variety of vectors suitable to this aspect of the invention, including constitutive and inducible expression vectors for use in prokaryotic and eukaryotic hosts, are well known and employed routinely by those of skill in the art.

The engineered host cells can be cultured in conventional nutrient media, which may be modified as appropriate for, *inter alia*, activating promoters, selecting transformants or amplifying genes. Culture conditions, such as temperature, pH and the like, previously used with the host cell selected for expression generally will be suitable for expression of polypeptides of the present invention as will be apparent to those of skill in the art.

A great variety of expression vectors can be used to express a polypeptide of the invention. Such vectors include chromosomal, episomal and virus-derived vectors *e.g.*, vectors derived from bacterial plasmids, from bacteriophage, from yeast episomes, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids, all may be used for expression in accordance with this aspect of the present invention. Generally, any vector suitable to maintain, propagate or express polynucleotides to express a polypeptide in a host may be used for expression in this regard.

The appropriate DNA sequence may be inserted into the vector by any of a variety of well-known and routine techniques. In general, a DNA sequence for expression is joined to an expression vector by cleaving the DNA sequence and the expression vector with one or more restriction endonucleases and then joining the restriction fragments together using T4 DNA ligase. Procedures for restriction and ligation that can be used to this end are well known and routine to those of skill. Suitable procedures in this regard, and for constructing expression vectors using alternative techniques, which also are well known and routine to those skill, are set forth in great detail in Sambrook *et al.* cited elsewhere herein.

The DNA sequence in the expression vector is operatively linked to appropriate expression control sequence(s), including, for instance, a promoter to direct mRNA transcription. Representatives of such promoters include the phage lambda PL promoter, the *E. coli* lac, trp and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name just a few of the well-known promoters. It will be understood that numerous promoters not mentioned are suitable for use in this aspect of the invention are well known and readily may be employed by those of skill in the manner illustrated by the discussion and the examples herein.

In general, expression constructs will contain sites for transcription initiation and termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs will include a translation initiating AUG at the beginning and a termination codon appropriately positioned at the end of the polypeptide to be translated.

In addition, the constructs may contain control regions that regulate as well as engender expression. Generally, in accordance with many commonly practiced procedures, such regions will operate by controlling transcription, such as repressor binding sites and enhancers, among others.

Vectors for propagation and expression generally will include selectable markers. Such markers also may be suitable for amplification or the vectors may contain additional markers for this purpose. In this regard, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells. Preferred markers include dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, and tetracycline or ampicillin resistance genes for culturing *E. coli* and other bacteria.

The vector containing the appropriate DNA sequence as described elsewhere herein, as well as an appropriate promoter, and other appropriate control sequences, may be introduced into an appropriate host using a variety of well known techniques suitable to expression therein of a desired polypeptide. Representative examples of appropriate hosts include bacterial cells, such as *E. coli*, Streptomyces and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS and Bowes melanoma cells; and plant cells. Hosts for of a great variety of expression constructs are well known, and those of skill will be enabled by the present disclosure readily to select a host for expressing a polypeptides in accordance with this aspect of the present invention.

More particularly, the present invention also includes recombinant constructs, such as expression constructs, comprising one or more of the sequences described above. The constructs comprise a vector, such as a plasmid or viral vector, into which such a sequence of the invention has been inserted. The sequence may be inserted in a forward or reverse orientation. In certain preferred embodiments in this regard, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and there are many commercially available vectors suitable for use in the present invention.

The following vectors, which are commercially available, are provided by way of example. Among vectors preferred for use in bacteria are pQE70, pQE60 and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. These vectors are listed solely by way of illustration of the many commercially available and well known vectors that are available to those of skill in the art for use in accordance with this aspect of the present invention. It will be appreciated that any other plasmid or vector suitable for, for example, introduction, maintenance, propagation or expression of a polynucleotide or polypeptide of the invention in a host may be used in this aspect of the invention.

Promoter regions can be selected from any desired gene using vectors that contain a reporter transcription unit lacking a promoter region, such as a chloramphenicol acetyl transferase ("cat") transcription unit, downstream of restriction site or sites for introducing a candidate promoter fragment; *i.e.*, a fragment that may contain a promoter. As is well known, introduction into the vector of a promoter-containing fragment at the restriction site upstream of the cat gene engenders production

of CAT activity, which can be detected by standard CAT assays. Vectors suitable to this end are well known and readily available. Two such vectors are pKK232-8 and pCM7. Thus, promoters for expression of polynucleotides of the present invention include not only well known and readily available promoters, but also promoters that readily may be obtained by the foregoing technique, using a reporter gene.

Among known bacterial promoters suitable for expression of polynucleotides and polypeptides in accordance with the present invention are the *E. coli* lacI and lacZ and promoters, the T3 and T7 promoters, the T5 tac promoter, the lambda PR, PL promoters and the trp promoter. Among known eukaryotic promoters suitable in this regard are the CMV immediate early promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters, the promoters of retroviral LTRs, such as those of the Rous sarcoma virus ("RSV"), and metallothionein promoters, such as the mouse metallothionein-I promoter.

Selection of appropriate vectors and promoters for expression in a host cell is a well known procedure and the requisite techniques for expression vector construction, introduction of the vector into the host and expression in the host are routine skills in the art.

The present invention also relates to host cells containing the above-described constructs discussed above. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis *et al.* Basic Methods In Molecular Biology, (1986).

Constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, the polypeptides of the invention can be synthetically produced by conventional peptide synthesizers.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).

Generally, recombinant expression vectors will include origins of replication, a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence, and a selectable marker to permit isolation of vector containing cells after exposure to the vector.

Transcription of the DNA encoding the polypeptides of the present invention by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act to increase transcriptional activity of a promoter in a given host cell-type. Examples of enhancers include the SV40 enhancer, which is located on the late side of the replication origin at bp 100 to 270, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Polynucleotides of the invention, encoding the heterologous structural sequence of a polypeptide of the invention generally will be inserted into the vector using standard techniques so that it is operably linked to the promoter for expression. The polynucleotide will be positioned so that the transcription start site is located appropriately 5' to a ribosome binding site. The ribosome binding site will be 5' to the

AUG that initiates translation of the polypeptide to be expressed. Generally, there will be no other open reading frames that begin with an initiation codon, usually AUG, and lie between the ribosome binding site and the initiating AUG. Also, generally, there will be a translation stop codon at the end of the polypeptide and there will be a polyadenylation signal and a transcription termination signal appropriately disposed at the 3' end of the transcribed region.

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide. The signals may be endogenous to the polypeptide or they may be heterologous signals.

The polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals but also additional heterologous functional regions. Thus, for instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence in the host cell, during purification or during subsequent handling and storage. Also, region also may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability and to facilitate purification, among others, are familiar and routine techniques in the art.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, where the selected promoter is inducible it is induced by appropriate means (*e.g.*, temperature shift or exposure to chemical inducer) and cells are cultured for an additional period. Cells typically then are harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, such methods are well know to those skilled in the art.

Various mammalian cell culture systems can be employed for expression, as well. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblast, described in Gluzman *et al.*, *Cell, 23*: 175 (1981). Other cell lines capable of expressing a compatible vector include for example, the C127, 3T3, CHO, HeLa, human kidney 293 and BHK cell lines.

The AIM-I polypeptide can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification. Well known techniques for refolding protein may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

AIM-I polynucleotides and polypeptides may be used in accordance with the present invention for a variety of applications, particularly those that make use of the

chemical and biological properties AIM-I. Among these are applications in autoimmune disease and aberrant cellular proliferation. Additional applications relate to diagnosis and to treatment of disorders of cells, tissues and organisms. These aspects of the invention are illustrated further by the following discussion.

This invention is also related to the use of the AIM-I polynucleotides to detect complementary polynucleotides such as, for example, as a diagnostic reagent. Detection of a mutated form of AIM-I associated with a dysfunction will provide a diagnostic tool that can add or define a diagnosis of a disease or susceptibility to diseases which results from under-expression over-expression or altered expression of AIM-I, such as, for example, autoimmune diseases.

Individuals carrying mutations in the human AIM-I gene may be detected at the DNA level by a variety of techniques. Nucleic acids for diagnosis may be obtained from a patient's cells, such as from blood, urine, saliva, tissue biopsy and autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR prior to analysis. PCR (Saiki *et al.*, *Nature*, *324*:163-166, (1986). RNA or cDNA may also be used in the same ways. As an example, PCR primers complementary to the nucleic acid encoding AIM-I can be used to identify and analyze AIM-I expression and mutations. For example, deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to radiolabeled AIM-I RNA or alternatively, radiolabeled AIM-I antisense DNA sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase A digestion or by differences in melting temperatures.

Sequence differences between a reference gene and genes having mutations also may be revealed by direct DNA sequencing. In addition, cloned DNA segments may be employed as probes to detect specific DNA segments. The sensitivity of such methods can be greatly enhanced by appropriate use of PCR or another amplification

method. For example, a sequencing primer is used with double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures with radiolabeled nucleotide or by automatic sequencing procedures with fluorescent-tags.

Genetic testing based on DNA sequence differences may be achieved by detection of alteration in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents. Small sequence deletions and insertions can be visualized by high resolution gel electrophoresis. DNA fragments of different sequences may be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures (see, *e.g.*, Myers *et al.*, Science, 230: 1242, 1985).

Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (*e.g.*, Cotton *et al.*, *Proc. Natl. Acad. Sci., USA, 85*:4397-4401, 1985).

Thus, the detection of a specific DNA sequence may be achieved by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing or the use of restriction enzymes, (*e.g.*, restriction fragment length polymorphisms ("RFLP") and Southern blotting of genomic DNA.

In addition to more conventional gel-electrophoresis and DNA sequencing, mutations also can be detected by *in situ* analysis.

The sequences of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. Moreover, there is a current need for identifying particular sites on the chromosome. Few chromosome marking

reagents based on actual sequence data (repeat polymorphisms) are presently available for marking chromosomal location. The mapping of DNAs to chromosomes according to the present invention is an important first step in correlating those sequences with genes associated with disease.

In certain preferred embodiments in this regard, the cDNA herein disclosed is used to clone genomic DNA of an AIM-I gene. This can be accomplished using a variety of well known techniques and libraries, which generally are available commercially. The genomic DNA the is used for *in situ* chromosome mapping using well known techniques for this purpose. Typically, in accordance with routine procedures for chromosome mapping, some trial and error may be necessary to identify a genomic probe that gives a good *in situ* hybridization signal.

In some cases, in addition, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the cDNA. Computer analysis of the 3' untranslated region of the gene is used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the primer will yield an amplified fragment.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular DNA to a particular chromosome. Using the present invention with the same oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes or pools of large genomic clones in an analogous manner. Other mapping strategies that can similarly be used to map to its chromosome include *in situ* hybridization, prescreening with labeled flow-sorted chromosomes and preselection by hybridization to construct chromosome specific-cDNA libraries.

Fluorescence *in situ* hybridization ("FISH") of a cDNA clone to a metaphase chromosomal spread can be used to provide a precise chromosomal location in one step. This technique can be used with cDNA as short as 50 or 60. For a review of this technique, see Verma *et al.*, Human Chromosomes: A Manual Of Basic Techniques, Pergamon Press, New York (1988).

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in V. McKusick, Mendelian Inheritance In Man, available on line through Johns Hopkins University, Welch Medical Library. The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

Next, it is necessary to determine the differences in the cDNA or genomic sequence between affected and unaffected individuals. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

With current resolution of physical mapping and genetic mapping techniques, a cDNA precisely localized to a chromosomal region associated with the disease could be one of between 50 and 500 potential causative genes. (This assumes 1 megabase mapping resolution and one gene per 20 kb).

The present invention also relates to a diagnostic assays such as quantitative and diagnostic assays for detecting levels of AIM-I protein in cells and tissues, including determination of normal and abnormal levels. Thus, for instance, a diagnostic assay in accordance with the invention for detecting over-expression of AIM-I protein compared to normal control tissue samples may be used to detect the presence of aberrant cellular proliferation such as cancer, for example. Assay techniques that can be used to

determine levels of a protein, such as an AIM-I protein of the present invention, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays. Among these ELISAs frequently are preferred. An ELISA assay initially comprises preparing an antibody specific to AIM-I, preferably a monoclonal antibody. In addition a reporter antibody generally is prepared which binds to the monoclonal antibody. The reporter antibody is attached a detectable reagent such as radioactive, fluorescent or enzymatic reagent, in this example horseradish peroxidase enzyme.

To carry out an ELISA a sample is removed from a host and incubated on a solid support, *e.g.* a polystyrene dish, that binds the proteins in the sample. Any free protein binding sites on the dish are then covered by incubating with a non-specific protein such as bovine serum albumin. Next, the monoclonal antibody is incubated in the dish during which time the monoclonal antibodies attach to any AIM-I proteins attached to the polystyrene dish. Unbound monoclonal antibody is washed out with buffer. The reporter antibody linked to horseradish peroxidase is placed in the dish resulting in binding of the reporter antibody to any monoclonal antibody bound to AIM-I. Unattached reporter antibody is then washed out. Reagents for peroxidase activity, including a colorimetric substrate are then added to the dish. Immobilized peroxidase, linked to AIM-I through the primary and secondary antibodies, produces a colored reaction product. The amount of color developed in a given time period indicates the amount of AIM-I protein present in the sample. Quantitative results typically are obtained by reference to a standard curve.

A competition assay may be employed wherein antibodies specific to AIM-I attached to a solid support and labeled AIM-I and a sample derived from the host are passed over the solid support and the amount of label detected attached to the solid support can be correlated to a quantity of AIM-I in the sample.

The polypeptides, their fragments or other derivatives, or analogs thereof, or cells expressing them can be used as an immunogen to produce antibodies thereto. These antibodies can be, for example, polyclonal or monoclonal antibodies. The present invention also includes chimeric, single chain, and humanized antibodies, as well as Fab fragments, or the product of an Fab expression library. Various procedures known in the art may be used for the production of such antibodies and fragments.

Antibodies generated against the polypeptides corresponding to a sequence of the present invention can be obtained by direct injection of the polypeptides into an animal or by administering the polypeptides to an animal, preferably a nonhuman. The antibody so obtained will then bind the polypeptides itself. In this manner, even a sequence encoding only a fragment of the polypeptides can be used to generate antibodies binding the whole native polypeptides. Such antibodies can then be used to isolate the polypeptide from tissue expressing that polypeptide.

For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., *Nature, 256*:495-497 (1975), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, *Immunology Today*, *4*:72 (1983) and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole *et al.*, pg. 77-96 in Monoclonal Antibodies And Cancer Therapy, Alan R. Liss, Inc. (1985).

Techniques described for the production of single chain antibodies (U.S. Patent No. 4,946,778) can be adapted to produce single chain antibodies to immunogenic polypeptide products of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies to immunogenic polypeptide products of this invention.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or purify the polypeptide of the present invention by attachment of the antibody to a solid support for isolation and/or purification by affinity chromatography. These antibodies may be employed to treat auto-immune diseases by preventing the ligand from binding its receptor.

Thus, among others, the AIM-I of the present invention may be employed to treat lymphoproliferative disease which results in lymphadenopathy, the AIM-I mediates apoptosis by stimulating clonal deletion of T-cells and may therefore, by employed to treat autoimmune disease, to stimulate peripheral tolerance and cytotoxic T-cell mediated apoptosis. The AIM-I may also be employed as a research tool in elucidating the biology of autoimmune disorders including systemic lupus erythematosus, immunoproliferative disease lymphadenopathy (IPL), angioimmunoproliferative lymphadenopathy (AIL), rheumatoid arthritis, diabetes, and multiple sclerosis, and to treat graft versus host disease.

The AIM-I may of the present invention may also be employed to inhibit neoplasia, such as tumor cell growth. The AIM-I polypeptide may be responsible for tumor destruction through apoptosis and cytotoxicity to certain cells. AIM-I may also be employed to treat diseases which require growth promotion activity, for example, restenosis, since AIM-I has proliferation effects on cells of endothelial origin. AIM-I may, therefore, also be employed to regulate hematopoiesis in endothelial cell development.

The polynucleotide encoding the AIM-I may be employed as a diagnostic marker for determining expression of the polypeptide of the present invention since the gene is found in many tumor cell lines including pancreatic tumor, testes tumor, endometrial tumor and T-cell lymphoma.

This invention also provides a method for identification of molecules, such as receptor molecules, that bind AIM-I. Genes encoding proteins that bind AIM-I, such as receptor proteins, can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting. Such methods are described in many laboratory manuals such as, for instance, Coligan *et al.*, Current Protocols in Immunology 1(2): Chapter 5 (1991).

For instance, expression cloning may be employed for this purpose. To this end polyadenylated RNA is prepared from a cell responsive to AIM-I, a cDNA library is created from this RNA, the library is divided into pools and the pools are transfected individually into cells that are not responsive to AIM-I. The transfected cells then are exposed to labeled AIM-I. (AIM-I can be labeled by a variety of well-known techniques including standard methods of radio-iodination or inclusion of a recognition site for a site-specific protein kinase.) Following exposure, the cells are fixed and binding of AIM-I is determined. These procedures conveniently are carried out on glass slides.

Pools are identified of cDNA that produced AIM-I-binding cells. Sub-pools are prepared from these positives, transfected into host cells and screened as described above. Using an iterative sub-pooling and re-screening process, one or more single clones that encode the putative binding molecule, such as a receptor molecule, can be isolated.

Alternatively a labeled ligand can be photoaffinity linked to a cell extract, such as a membrane or a membrane extract, prepared from cells that express a molecule that it binds, such as a receptor molecule. Cross-linked material is resolved by polyacrylamide gel electrophoresis ("PAGE") and exposed to X-ray film. The labeled complex containing the ligand-receptor can be excised, resolved into peptide fragments, and subjected to protein microsequencing. The amino acid sequence obtained from microsequencing can be used to design unique or degenerate oligonucleotide probes to screen cDNA libraries to identify genes encoding the putative receptor molecule.

Polypeptides of the invention also can be used to assess AIM-I binding capacity of AIM-I binding molecules, such as receptor molecules, in cells or in cell-free preparations.

The invention also provides a method of screening compounds to identify those which enhance or block the action of AIM-I on cells, such as its interaction with AIM-I-binding molecules such as receptor molecules. An agonist is a compound which increases the natural biological functions of AIM-I or which functions in a manner similar to AIM-I, while antagonists decrease or eliminate such functions.

For example, a cellular compartment, such as a membrane or a preparation thereof, such as a membrane-preparation, may be prepared from a cell that expresses a molecule that binds AIM-I, such as a molecule of a signaling or regulatory pathway modulated by AIM-I. The preparation is incubated with labeled AIM-I in the absence or the presence of a candidate molecule which may be an AIM-I agonist or antagonist. The ability of the candidate molecule to bind the binding molecule is reflected in decreased binding of the labeled ligand. Molecules which bind gratuitously, *i.e.*, without inducing the effects of AIM-I on binding the AIM-I binding molecule, are most likely to be good antagonists. Molecules that bind well and elicit effects that are the same as or closely related to AIM-I, are good agonists.

AIM-I-like effects of potential agonists and antagonists may by measured, for instance, by determining activity of a second messenger system following interaction of the candidate molecule with a cell or appropriate cell preparation, and comparing the effect with that of AIM-I or molecules that elicit the same effects as AIM-I. Second messenger systems that may be useful in this regard include but are not limited to AMP guanylate cyclase, ion channel or phosphoinositide hydrolysis second messenger systems.

Another example of an assay for AIM-I antagonists is a competitive assay that combines AIM-I and a potential antagonist with membrane-bound AIM-I receptor molecules or recombinant AIM-I receptor molecules under appropriate conditions for a competitive inhibition assay. AIM-I can be labeled, such as by radioactivity, such that the number of AIM-I molecules bound to a receptor molecule can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to a polypeptide of the invention, and thereby inhibit or extinguish its activity. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as a receptor molecule, without inducing AIM-I-induced activities, thereby preventing the action of AIM-I by excluding AIM-I from binding.

Other potential antagonists include antisense molecules. Antisense technology can be used to control gene expression through antisense DNA or RNA or through triple-helix formation. Antisense techniques are discussed, for example, in Okano, *J. Neurochem*, *56*:560, 1991; Oligodeoxynucleotides As Antisense Inhibitors Of Gene Expression, CRC Press, Boca Raton, FL (1988). Triple helix formation is discussed in, for instance Lee *et al.*, *Nucleic Acids Research*, *6*:3073, 1979; Cooney *et al.*, *Science*, *241*:456 1988; and Dervan *et al.*, *Science*, *251*:1360 1991. The methods are based on binding of a polynucleotide to a complementary DNA or RNA. For example, the 5' coding portion of a polynucleotide that encodes the mature polypeptide of the present invention may be used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription thereby preventing transcription and the production of AIM-I. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into AIM-I polypeptide. The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of AIM-I.

The antagonists may be employed in a composition with a pharmaceutically acceptable carrier, *e.g.*, as hereinafter described.

The antagonists may be employed for instance to treat cachexia which is a lipid clearing defect resulting from a systemic deficiency of lipoprotein lipase, which is suppressed by AIM-I. The AIM-I antagonists may also be employed to treat cerebral malaria in which AIM-I appears to play a pathogenic role. The antagonists may also be employed to treat rheumatoid arthritis by inhibiting AIM-I induced production of inflammatory cytokines, such as IL1 in the synovial cells. When treating arthritis, AIM-I is preferably injected intra-articularly.

The AIM-I antagonists may also be employed to prevent graft-host rejection by preventing the stimulation of the immune system in the presence of a graft.

The AIM-I antagonists may also be employed to inhibit bone resorption and, therefore, to treat and/or prevent osteoporosis.

The antagonists may also be employed as anti-inflammatory agents, and to treat endotoxic shock. This critical condition results from an exaggerated response to bacterial and other types of infection.

The invention also relates to compositions comprising the polynucleotide or the polypeptides discussed above or the agonists or antagonists. Thus, the polypeptides of the present invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to a subject. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of a polypeptide of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation should suit the mode of administration.

The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, reflecting approval by the agency of the manufacture, use or sale of the product for human administration.

Polypeptides and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes among others.

The pharmaceutical compositions generally are administered in an amount effective for treatment or prophylaxis of a specific indication or indications. In general, the compositions are administered in an amount of at least about 10 $\mu$g/kg body weight. In most cases they will be administered in an amount not in excess of about 8 mg/kg body weight per day. Preferably, in most cases, dose is from about 10 $\mu$g/kg to about 1 mg/kg body weight, daily. It will be appreciated that optimum dosage will be determined by standard methods for each treatment modality and indication, taking into account the indication, its severity, route of administration, complicating conditions and the like.

The AIM-1 polynucleotides, polypeptides, agonists and antagonists that are polypeptides may be employed in accordance with the present invention by expression of such polypeptides *in vivo*, in treatment modalities often referred to as "gene therapy."

Thus, for example, cells from a patient may be engineered with a polynucleotide, such as a DNA or RNA, encoding a polypeptide *ex vivo*, and the engineered cells then can be provided to a patient to be treated with the polypeptide. For example, cells may be engineered *ex vivo* by the use of a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention. Such methods are well-known in the art and their use in the present invention will be apparent from the teachings herein.

Similarly, cells may be engineered *in vivo* for expression of a polypeptide *in vivo* by procedures known in the art. For example, a polynucleotide of the invention may be engineered for expression in a replication defective retroviral vector, as discussed above. The retroviral expression construct then may be isolated and introduced into a packaging cell is transduced with a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a patient for engineering cells *in vivo* and expression of the polypeptide *in vivo*. These and other methods for administering a polypeptide of the present invention by such method should be apparent to those skilled in the art from the teachings of the present invention.

Retroviruses from which the retroviral plasmid vectors herein above mentioned may be derived include, but are not limited to, Moloney Murine Leukemia Virus, spleen necrosis virus, retroviruses such as Rous Sarcoma Virus, Harvey Sarcoma Virus, avian leukosis virus, gibbon ape leukemia virus, human immunodeficiency virus, adenovirus, Myeloproliferative Sarcoma Virus, and mammary tumor virus. In one embodiment, the retroviral plasmid vector is derived from Moloney Murine Leukemia Virus.

Such vectors well include one or more promoters for expressing the polypeptide. Suitable promoters which may be employed include, but are not limited to, the retroviral LTR; the SV40 promoter; and the human cytomegalovirus (CMV) promoter described in Miller *et al.*, *Biotechniques*, 7:980-990 (1989), or any other promoter (*e.g.*, cellular

promoters such as eukaryotic cellular promoters including, but not limited to, the histone, RNA polymerase III, and ß-actin promoters). Other viral promoters which may be employed include, but are not limited to, adenovirus promoters, thymidine kinase (TK) promoters, and B19 parvovirus promoters. The selection of a suitable promoter will be apparent to those skilled in the art from the teachings contained herein.

The nucleic acid sequence encoding the polypeptide of the present invention will be placed under the control of a suitable promoter. Suitable promoters which may be employed include, but are not limited to, adenoviral promoters, such as the adenoviral major late promoter; or heterologous promoters, such as the cytomegalovirus (CMV) promoter; the respiratory syncytial virus (RSV) promoter; inducible promoters, such as the MMT promoter, the metallothionein promoter; heat shock promoters; the albumin promoter; the ApoAI promoter; human globin promoters; viral thymidine kinase promoters, such as the Herpes Simplex thymidine kinase promoter; retroviral LTRs (including the modified retroviral LTRs herein above described); the ß-actin promoter; and human growth hormone promoters. The promoter also may be the native promoter which controls the gene encoding the polypeptide.

The retroviral plasmid vector is employed to transduce packaging cell lines to form producer cell lines. Examples of packaging cells which may be transfected include, but are not limited to, the PE501, PA317, Y-2, Y-AM, PA12, T19-14X, VT-19-17-H2, YCRE, YCRIP, GP+E-86, GP+envAm12, and DAN cell lines as described in Miller, A., *Human Gene Therapy*, *1*:5-14 (1990). The vector may be transduced into the packaging cells through any means known in the art. Such means include, but are not limited to, electroporation, the use of liposomes, and $CaPO_4$ precipitation. In one alternative, the retroviral plasmid vector may be encapsulated into a liposome, or coupled to a lipid, and then administered to a host.

The producer cell line will generate infectious retroviral vector particles, which include the nucleic acid sequence(s) encoding the polypeptides. Such retroviral vector

particles then may be employed to transduce eukaryotic cells, either *in vitro* or *in vivo*. The transduced eukaryotic cells will express the nucleic acid sequence(s) encoding the polypeptide. Eukaryotic cells which may be transduced include, but are not limited to, embryonic stem cells, embryonic carcinoma cells, as well as hematopoietic stem cells, hepatocytes, fibroblasts, myoblasts, keratinocytes, endothelial cells, and bronchial epithelial cells.

The present invention is further described by the following examples. The examples are provided solely to illustrate the invention by reference to specific embodiments. These exemplification's, while illustrating certain specific aspects of the invention, do not portray the limitations or circumscribe the scope of the disclosed invention.

Certain terms used herein are explained in the foregoing glossary. All examples were carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described in detail. Routine molecular biology techniques of the following examples can be carried out as described in standard laboratory manuals, such as Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989), herein referred to as "Sambrook." All parts or amounts set out in the following examples are by weight, unless otherwise specified.

Unless otherwise stated size separation of fragments in the examples below was carried out using standard techniques of agarose and polyacrylamide gel electrophoresis ("PAGE") in Sambrook and numerous other references such as, for instance, by Goeddel *et al.*, *Nucleic Acids Res.* 8:4057 (1980).

Unless described otherwise, ligations were accomplished using standard buffers, incubation temperatures and times, approximately equimolar amounts of the DNA

fragments to be ligated and approximately 10 units of T4 DNA ligase ("ligase") per 0.5 μg of DNA.

## Example 1

## Expression and Purification of Human AIM-I Using Bacteria

The DNA sequence encoding human AIM-I in the deposited polynucleotide was amplified using PCR oligonucleotide primers specific to the amino acid carboxyl terminal sequence of the human AIM-I protein and to vector sequences 3' to the gene. Additional nucleotides containing restriction sites to facilitate cloning were added to the 5' and 3' sequences respectively.

The 5' oligonucleotide primer had the sequence 5' GCG GCG GGA TCC ATG GCT ATG ATG GAG GTC CAG 3' containing the underlined BamHI restriction site, which encodes a start AUG, followed by 18 nucleotides of the human AIM-I coding sequence set out in Figure 1.

The 3' primer had the sequence 5'CGC GCG TCT AGA GCT TAG GCA ACT AAA AAG GCC 3' containing the underlined XbaI restriction site followed by 21 nucleotides complementary to the last 21 nucleotides of the AIM-I coding sequence set out in Figure 1, including the stop codon.

The restrictions sites were convenient to restriction enzyme sites in the bacterial expression vectors pQE9 which were used for bacterial expression in these examples. (Qiagen, Inc., Chatsworth, CA). pQE9 encodes ampicillin antibiotic resistance ("Amp'") and contains a bacterial origin of replication ("ori"), an IPTG inducible promoter, a ribosome binding site ("RBS"), a 6-His tag and restriction enzyme sites.

The amplified human AIM-I DNA and the vector pQE9 both were digested with BamHI and XbaI and the digested DNAs then were ligated together. Insertion of the

AIM-I DNA into the restricted vector placed the AIM-I coding region downstream of and operably linked to the vector's IPTG-inducible promoter and in-frame with an initiating AUG appropriately positioned for translation of AIM-I.

The ligation mixture was transformed into competent *E. coli* cells using standard procedures. Such procedures are described in Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2nd Ed.: Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989). *E. coli* strain M15/rep4, containing multiple copies of the plasmid pREP4, which expresses lac repressor and confers kanamycin resistance ("Kan$^r$"), was used in carrying out the illustrative example described here. This strain, which is only one of many that are suitable for expressing AIM-I, is available commercially from Qiagen.

Transformants were identified by their ability to grow on LB plates in the presence of ampicillin. Plasmid DNA was isolated from resistant colonies and the identity of the cloned DNA was confirmed by restriction analysis.

Clones containing the desired constructs were grown overnight ("O/N") in liquid culture in LB media supplemented with both ampicillin (100 μg/ml) and kanamycin (25 μg/ml).

The O/N culture was used to inoculate a large culture, at a dilution of approximately 1:100 to 1:250. The cells were grown to an optical density at 600 nm ("$OD_{600}$") of between 0.4 and 0.6. Isopropyl-B-D-thiogalactopyranoside ("IPTG") was then added to a final concentration of 1 mM to induce transcription from lac repressor sensitive promoters, by inactivating the lacI repressor. Cells subsequently were incubated further for 3 to 4 hours. Cells then were harvested by centrifugation and disrupted, by standard methods. Inclusion bodies were purified from the disrupted cells using routine collection techniques, and protein was solubilized from the inclusion bodies into 8M urea. The 8M urea solution containing the solubilized protein was

passed over a PD-10 column in 2X phosphate buffered saline ("PBS"), thereby removing the urea, exchanging the buffer and refolding the protein. The protein was purified by a further step of chromatography to remove endotoxin. Then, it was sterile filtered. The sterile filtered protein preparation was stored in 2X PBS at a concentration of 95 micrograms per mL.

Analysis of the preparation by standard methods of polyacrylamide gel electrophoresis revealed that the preparation contained above 90% monomer AIM-I having the expected molecular weight of, approximately, 31 kDa.

## Example 2

## Cloning and Expression of Human AIM-I in a Baculovirus Expression System

The cDNA sequence encoding the full length human AIM-I protein, in the deposited clone is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the gene:

The 5' primer has the sequence 5' CCG CGC GGA TCC ATC ATG GCT ATG ATG GAG GTC C 3' containing the underlined restriction enzyme site followed by 22 bases of the sequence of AIM-I of Figure 1. Inserted into an expression vector, as described below, the 5' end of the amplified fragment encoding human AIM-I provides an efficient signal peptide. An efficient signal for initiation of translation in eukaryotic cells, as described by Kozak, M., *J. Mol. Biol.*, *196*:947-950 (1987) is appropriately located in the vector portion of the construct.

The 3' primer has the sequence 5' CGC GCG TCT AGA GCT TAG CCA ACT AAA AAG GCC 3' containing the underlined XbaI restriction followed by nucleotides complementary to the last 21 nucleotides of the AIM-I coding sequence set out in Figure 1, including the stop codon.

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with BamH1 and Asp718 and again is purified on a 1% agarose gel. This fragment is designated herein F2.

The vector pA2 is used to express the AIM-I protein in the baculovirus expression system, using standard methods, such as those described in Summers *et al.*, A Manual Of Methods For Baculovirus Vectors And Insect Cell Culture Procedures, Texas Agricultural Experimental Station Bulletin No. 1555 (1987). This expression vector contains the strong polyhedrin promoter of the Autographa californica nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites. The signal peptide of AcMNPV gp67, including the N-terminal methionine, is located just upstream of a BamH1 site. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For an easy selection of recombinant virus the beta-galactosidase gene from *E. coli* is inserted in the same orientation as the polyhedrin promoter and is followed by the polyadenylation signal of the polyhedrin gene. The polyhedrin sequences are flanked at both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate viable virus that express the cloned polynucleotide.

Many other baculovirus vectors could be used in place of pA2, such as pAc373, pVL941 and pAcIM1 provided, as those of skill readily will appreciate, that construction provides appropriately located signals for transcription, translation, trafficking and the like, such as an in-frame AUG and a signal peptide, as required. Such vectors are described in Luckow *et al.*, *Virology*, *170*:31-39, among others.

The plasmid is digested with the restriction enzymes BamHI and XbaI and then is dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially

available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.). This vector DNA is designated herein "V2".

Fragment F2 and the dephosphorylated plasmid V2 are ligated together with T4 DNA ligase. *E. coli* HB101 cells are transformed with ligation mix and spread on culture plates. Bacteria are identified that contain the plasmid with the human AIM-I gene by digesting DNA from individual colonies using BamHI and XbaI and then analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing. This plasmid is designated herein pBacAIM-I.

5 μg of the plasmid pBacAIM-I is co-transfected with 1.0 μg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA.), using the lipofection method described by Felgner *et al.*, *Proc. Natl. Acad. Sci. USA*, *84*:7413-7417 (1987). 1μg of BaculoGold™ virus DNA and 5 μg of the plasmid pBacAIM-I are mixed in a sterile well of a microtiter plate containing 50 μl of serum free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards 10 μl Lipofectin plus 90 μl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is rocked back and forth to mix the newly added solution. The plate is then incubated for 5 hours at 27°C. After 5 hours the transfection solution is removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. The plate is put back into an incubator and cultivation is continued at 27°C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, cited above. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of

a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10).

Four days after serial dilution, the virus is added to the cells. After appropriate incubation, blue stained plaques are picked with the tip of an Eppendorf pipette. The agar containing the recombinant viruses is then resuspended in an Eppendorf tube containing 200 $\mu$l of Grace's medium. The agar is removed by a brief centrifugation and the supernatant containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4°C. A clone containing properly inserted AIM-I is identified by DNA analysis including restriction mapping and sequencing. This is designated herein as V-AIM-I.

Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus V-AIM-I at a multiplicity of infection ("MOI") of about 2 (about 1 to about 3). Six hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Gaithersburg). 42 hours later, 5 $\mu$Ci of $^{35}$S-methionine and 5 $\mu$Ci $^{35}$S cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then they are harvested by centrifugation, lysed and the labeled proteins are visualized by SDS-PAGE and autoradiography.

## Example 3

## Expression of AIM-I in COS Cells

The expression plasmid, AIM-I HA, is made by cloning a cDNA encoding AIM-I into the expression vector pcDNAI/Amp (which can be obtained from Invitrogen, Inc.).

The expression vector pcDNAI/amp contains: (1) an *E. coli* origin of replication effective for propagation in *E. coli* and other prokaryotic cell; (2) an ampicillin resistance gene for selection of plasmid-containing prokaryotic cells; (3) an SV40 origin of replication for propagation in eukaryotic cells; (4) a CMV promoter, a polylinker, an SV40 intron, and a polyadenylation signal arranged so that a cDNA conveniently can be placed under expression control of the CMV promoter and operably linked to the SV40 intron and the polyadenylation signal by means of restriction sites in the polylinker.

A DNA fragment encoding the entire AIM-I precursor and a HA tag fused in frame to its 3' end is cloned into the polylinker region of the vector so that recombinant protein expression is directed by the CMV promoter. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein described by Wilson *et al.*, Cell 37: 767 (1984). The fusion of the HA tag to the target protein allows easy detection of the recombinant protein with an antibody that recognizes the HA epitope.

The plasmid construction strategy is as follows. The AIM-I cDNA of the deposited clone is amplified using primers that contained convenient restriction sites, much as described above regarding the construction of expression vectors for expression of AIM-I in *E. coli* and *S. furgiperda*.

To facilitate detection, purification and characterization of the expressed AIM-I, one of the primers contains a hemagglutinin tag ("HA tag") as described above.

Suitable primers include that following, which are used in this example.

The 5' primer, containing the underlined restriction enzyme site, an AUG start codon and 22 codons thereafter, forming the hexapeptide haemagglutinin tag, has the sequence: 5' CCG CGC GGA TCC ATC ATG GCT ATG ATG GAG GTC C 3' The 3' primer, containing the underlined XbaI site and 21 nucleotides of 3' coding sequence

(at the 3' end) has the sequence: 5' CGC GCG TCT AGA GCT TAG CCA ACT AAA AAG GCC 3'

The PCR amplified DNA fragment and the vector, pcDNAI/Amp, are digested with and then ligated. The ligation mixture is transformed into *E. coli* strain SURE (available from Stratagene Cloning Systems, 11099 North Torrey Pines Road, La Jolla, CA 92037) the transformed culture is plated on ampicillin media plates which then are incubated to allow growth of ampicillin resistant colonies. Plasmid DNA is isolated from resistant colonies and examined by restriction analysis and gel sizing for the presence of the AIM-I-encoding fragment.

For expression of recombinant AIM-I, COS cells are transfected with an expression vector, as described above, using DEAE-DEXTRAN, as described, for instance, in Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Laboratory Press, Cold Spring Harbor, New York (1989). Cells are incubated under conditions for expression of AIM-I by the vector.

Expression of the AIM-I HA fusion protein is detected by radiolabelling and immunoprecipitation, using methods described in, for example Harlow *et al.*, Antibodies: A Laboratory Manual, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1988). To this end, two days after transfection, the cells are labeled by incubation in media containing $^{35}$S-cysteine for 8 hours. The cells and the media are collected, and the cells are washed and the lysed with detergent-containing RIPA buffer: 150 mM NaCl, 1% NP-40, 0.1% SDS, 1% NP-40, 0.5% DOC, 50 mM TRIS, pH 7.5, as described by Wilson *et al.* cited above. Proteins are precipitated from the cell lysate and from the culture media using an HA-specific monoclonal antibody. The precipitated proteins then are analyzed by SDS-PAGE gels and autoradiography. An expression product of the expected size is seen in the cell lysate, which is not seen in negative controls.

## Example 4

## Tissue Distribution of AIM-I Expression

Northern blot analysis is carried out to examine the levels of expression of AIM-I in human tissues, using methods described by, among others, Sambrook *et al*, cited above. Total cellular RNA samples are isolated with RNAzol™ B system (Biotecx Laboratories, Inc. 6023 South Loop East, Houston, TX 77033).

About 10$\mu$g of total RNA is isolated from tissue samples. The RNA is size resolved by electrophoresis through a 1% agarose gel under strongly denaturing conditions. RNA is blotted from the gel onto a nylon filter, and the filter then is prepared for hybridization to a detectably labeled polynucleotide probe.

As a probe to detect mRNA that encodes AIM-I, the antisense strand of the coding region of the cDNA insert in the deposited clone, is labeled to a high specific activity. The cDNA is labeled by primer extension, using the Prime-It kit, available from Stratagene. The reaction is carried out using 50 ng of the cDNA, following the standard reaction protocol as recommended by the supplier. The labeled polynucleotide is purified away from other labeled reaction components by column chromatography using a Select-G-50 column, obtained from 5-Prime - 3-Prime, Inc. of 5603 Arapahoe Road, Boulder, CO 80303.

The labeled probe is hybridized to the filter, at a concentration of 1,000,000 cpm/ml, in a small volume of 7% SDS, 0.5 M $NaPO_4$, pH 7.4 at 65°C, overnight. Thereafter the probe solution is drained and the filter is washed twice at room temperature and twice at 60°C with 0.5 x SSC, 0.1% SDS. The filter then is dried and exposed to film at -70°C overnight with an intensifying screen. Autoradiography shows that mRNA for AIM-I is abundant in human heart, bone marrow, $CD4^+$ and $CD19^+$ peripheral blood lymphocytes, and less so in lung and kidney tissue.

## Example 5

## Gene Therapeutic Expression of Human AIM-I

Fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature overnight. After 24 hours at room temperature, the flask is inverted - the chunks of tissue remain fixed to the bottom of the flask - and fresh media is added (*e.g.*, Ham's F12 media, with 10% FBS, penicillin and streptomycin). The tissue is then incubated at 37°C for approximately one week. At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerges. The monolayer is trypsinized and scaled into larger flasks.

A vector for gene therapy is digested with restriction enzymes for cloning a fragment to be expressed. The digested vector is treated with calf intestinal phosphatase to prevent self-ligation. The dephosphorylated, linear vector is fractionated on an agarose gel and purified.

AIM-I cDNA capable of expressing active AIM-I, is isolated. The ends of the fragment are modified, if necessary, for cloning into the vector. For instance, 5" overhanging may be treated with DNA polymerase to create blunt ends. 3' overhanging ends may be removed using S1 nuclease. Linkers may be ligated to blunt ends with T4 DNA ligase.

Equal quantities of the Moloney murine leukemia virus linear backbone and the AIM-I fragment are mixed together and joined using T4 DNA ligase. The ligation mixture is used to transform *E. coli* and the bacteria are then plated onto agar-containing

kanamycin. Kanamycin phenotype and restriction analysis confirm that the vector has the properly inserted gene.

Packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagle's Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The vector containing the AIM-I gene is introduced into the packaging cells by standard techniques. Infectious viral particles containing the AIM-I gene are collected from the packaging cells, which now are called producer cells.

Fresh media is added to the producer cells, and after an appropriate incubation period media is harvested from the plates of confluent producer cells. The media, containing the infectious viral particles, is filtered through a Millipore filter to remove detached producer cells. The filtered media then is used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the filtered media. Polybrene (Aldrich) may be included in the media to facilitate transduction. After appropriate incubation, the media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his, to select out transduced cells for expansion.

Engineered fibroblasts then may be injected into rats, either alone or after having been grown to confluence on microcarrier beads, such as cytodex 3 beads. The injected fibroblasts produce AIM-I product, and the biological actions of the protein are conveyed to the host.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Ruben, Steven M

(ii) TITLE OF INVENTION: Apoptosis Inducing Molecule I

(iii) NUMBER OF SEQUENCES: 10

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart & Olstein
(B) STREET: 6 Becker Farm Road
(C) CITY: Roseland
(D) STATE: NJ
(E) COUNTRY: US
(F) ZIP: 07068

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: PCT/US96/03773
(B) FILING DATE: 14-MAR-1996
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: MULLINS, J.G.
(B) REGISTRATION NUMBER: 33,073
(C) REFERENCE/DOCKET NUMBER: 325800-549

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (201) 994-1700
(B) TELEFAX: (201) 994-1744

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1643 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 52..894

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGCACGAGCG GCTGCCTGGC TGACTTACAG CAGTCAGACT CTGACAGGTT C ATG GCT 57
Met Ala
1

ATG ATG GAG GTC CAG GGG GGA CCC AGC CTG GGA CAG ACC TGC GTG CTG 105

```
Met Met Glu Val Gln Gly Gly Pro Ser Leu Gly Gln Thr Cys Val Leu
          5                  10                  15

ATC GTG ATC TTC ACA GTG CTC CTG CAG TCT CTC TGT GTG GCT GTA ACT      153

Ile Val Ile Phe Thr Val Leu Leu Gln Ser Leu Cys Val Ala Val Thr
     20                  25                  30

TAC GTG TAC TTT ACC AAC GAG CTG AAG CAG ATG CAG GAC AAG TAC TCC      201

Tyr Val Tyr Phe Thr Asn Glu Leu Lys Gln Met Gln Asp Lys Tyr Ser
 35                  40                  45                  50

AAA AGT GGC ATT GCT TGT TTC TTA AAA GAA GAT GAC AGT TAT TGG GAC      249

Lys Ser Gly Ile Ala Cys Phe Leu Lys Glu Asp Asp Ser Tyr Trp Asp
                 55                  60                  65

CCC AAT GAC GAA GAG AGT ATG AAC AGC CCC TGC TGG CAA GTC AAG TGG      297

Pro Asn Asp Glu Glu Ser Met Asn Ser Pro Cys Trp Gln Val Lys Trp
             70                  75                  80

CAA CTC CGT CAG CTC GTT AGA AAG ATG ATT TTG AGA ACC TCT GAG GAA      345

Gln Leu Arg Gln Leu Val Arg Lys Met Ile Leu Arg Thr Ser Glu Glu
         85                  90                  95

ACC ATT TCT ACA GTT CAA GAA AAG CAA CAA AAT ATT TCT CCC CTA GTG      393

Thr Ile Ser Thr Val Gln Glu Lys Gln Gln Asn Ile Ser Pro Leu Val
    100                 105                 110

AGA GAA AGA GGT CCT CAG AGA GTA GCA GCT CAC ATA ACT GGG ACC AGA      441

Arg Glu Arg Gly Pro Gln Arg Val Ala Ala His Ile Thr Gly Thr Arg
115                 120                 125                 130

GGA AGA AGC AAC ACA TTG TCT TCT CCA AAC TCC AAG AAT GAA AAG GCT      489

Gly Arg Ser Asn Thr Leu Ser Ser Pro Asn Ser Lys Asn Glu Lys Ala
                135                 140                 145

CTG GGC CGC AAA ATA AAC TCC TGG GAA TCA TCA AGG AGT GGG CAT TCA      537

Leu Gly Arg Lys Ile Asn Ser Trp Glu Ser Ser Arg Ser Gly His Ser
            150                 155                 160

TTC CTG AGC AAC TTG CAC TTG AGG AAT GGT GAA CTG GTC ATC CAT GAA      585

Phe Leu Ser Asn Leu His Leu Arg Asn Gly Glu Leu Val Ile His Glu
        165                 170                 175

AAA GGG TTT TAC TAC ATC TAT TCC CAA ACA TAC TTT CGA TTT CAG GAG      633

Lys Gly Phe Tyr Tyr Ile Tyr Ser Gln Thr Tyr Phe Arg Phe Gln Glu
    180                 185                 190

GAA ATA AAA GAA AAC ACA AAG AAC GAC AAA CAA ATG GTC CAA TAT ATT      681

Glu Ile Lys Glu Asn Thr Lys Asn Asp Lys Gln Met Val Gln Tyr Ile
195                 200                 205                 210

TAC AAA TAC ACA AGT TAT CCT GAC CCT ATA TTG TTG ATG AAA AGT GCT      729

Tyr Lys Tyr Thr Ser Tyr Pro Asp Pro Ile Leu Leu Met Lys Ser Ala
                215                 220                 225
```

```
AGA AAT AGT TGT TGG TCT AAA GAT GCA GAA TAT GGA CTC TAT TCC ATC        777

Arg Asn Ser Cys Trp Ser Lys Asp Ala Glu Tyr Gly Leu Tyr Ser Ile
        230                 235                 240

TAT CAA GGG GGA ATA TTT GAG CTT AAG GAA AAT GAC AGA ATT TTT GTT        825

Tyr Gln Gly Gly Ile Phe Glu Leu Lys Glu Asn Asp Arg Ile Phe Val
    245                 250                 255

TCT GTA ACA AAT GAG CAC TTG ATA GAC ATG GAC CAT GAA GCC AGT TTT        873

Ser Val Thr Asn Glu His Leu Ile Asp Met Asp His Glu Ala Ser Phe
260                 265                 270

TTC GGG GCC TTT TTA GTT GGC TAACTGACCT GGAAAGAAAA AGCAATAACC           924

Phe Gly Ala Phe Leu Val Gly
275                 280

TCAAAGTGAC TATTCAGTTT TCAGGATGAT ACACTATGAA GATGTTTCAA AAAATCTGAC      984

CAAAACAAAC AAACAGAAAA CAGAAAACAA AAAAACCTCT ATGCAATCTG AGTAGAGCAG     1044

CCACAACCAA AAAATTCTAC AACACACACT GTTCTGAAAG TGACTCACTT ATCCCAAGAA     1104

AATGAAATTG CTGAAAGATC TTTCAGGACT CTACCTCATA TCAGTTTGCT AGCAGAAATC     1164

TAGAAGACTG TCAGCTTCCA AACATTAATG CAATGGTTAA CATCTTCTGT CTTTATAATC     1224

TACTCCTTGT AAAGACTGTA GAAGAAAGCG CAACAATCCA TCTCTCAAGT AGTGTATCAC     1284

AGTAGTAGCC TCCAGGTTTC CTTAAGGGAC AACATCCTTA AGTCAAAAGA GAGAAGAGGC     1344

ACCACTAAAA GATCGCAGTT TGCCTGGTGC AGTGGCTCAC ACCTGTAATC CCAACATTTT     1404

GGGAACCCAA GGTGGGTAGA TCACGAGATC AAGAGATCAA GACCATAGTG ACCAACATAG     1464

TGAAACCCCA TCTCTACTGA AAGTGCAAAA ATTAGCTGGG TGTGTTGGCA CATGCCTGTA     1524

GTCCCAGCTA CTTGAGAGGC TGAGGCAGGA GAATCGTTTG AACCCGGGAG GCAGAGGTTG     1584

CAGTGTGGTG AGATCATGCC ACTACACTCC AGCCTGGCGA CAGAGCGAGA CTTGGTTTC      1643
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 281 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Ala Met Met Glu Val Gln Gly Gly Pro Ser Leu Gly Gln Thr Cys
 1               5                  10                  15

Val Leu Ile Val Ile Phe Thr Val Leu Leu Gln Ser Leu Cys Val Ala
            20                  25                  30

Val Thr Tyr Val Tyr Phe Thr Asn Glu Leu Lys Gln Met Gln Asp Lys
        35                  40                  45

Tyr Ser Lys Ser Gly Ile Ala Cys Phe Leu Lys Glu Asp Asp Ser Tyr
    50                  55                  60
```

```
Trp Asp Pro Asn Asp Glu Glu Ser Met Asn Ser Pro Cys Trp Gln Val
 65                  70                  75                  80

Lys Trp Gln Leu Arg Gln Leu Val Arg Lys Met Ile Leu Arg Thr Ser
                 85                  90                  95

Glu Glu Thr Ile Ser Thr Val Gln Glu Lys Gln Gln Asn Ile Ser Pro
            100                 105                 110

Leu Val Arg Glu Arg Gly Pro Gln Arg Val Ala Ala His Ile Thr Gly
        115                 120                 125

Thr Arg Gly Arg Ser Asn Thr Leu Ser Ser Pro Asn Ser Lys Asn Glu
    130                 135                 140

Lys Ala Leu Gly Arg Lys Ile Asn Ser Trp Glu Ser Ser Arg Ser Gly
145                 150                 155                 160

His Ser Phe Leu Ser Asn Leu His Leu Arg Asn Gly Glu Leu Val Ile
                165                 170                 175

His Glu Lys Gly Phe Tyr Tyr Ile Tyr Ser Gln Thr Tyr Phe Arg Phe
            180                 185                 190

Gln Glu Glu Ile Lys Glu Asn Thr Lys Asn Asp Lys Gln Met Val Gln
        195                 200                 205

Tyr Ile Tyr Lys Tyr Thr Ser Tyr Pro Asp Pro Ile Leu Leu Met Lys
    210                 215                 220

Ser Ala Arg Asn Ser Cys Trp Ser Lys Asp Ala Glu Tyr Gly Leu Tyr
225                 230                 235                 240

Ser Ile Tyr Gln Gly Gly Ile Phe Glu Leu Lys Glu Asn Asp Arg Ile
                245                 250                 255

Phe Val Ser Val Thr Asn Glu His Leu Ile Asp Met Asp His Glu Ala
            260                 265                 270

Ser Phe Phe Gly Ala Phe Leu Val Gly
        275                 280
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 267 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Val Asp Ser Ser Ala Ser Ser Pro Trp Ala Pro Pro Gly Thr Val Leu
1               5                   10                  15

Pro Cys Pro Thr Ser Val Pro Arg Arg Pro Gly Gln Arg Arg Pro Pro
            20                  25                  30

Pro Pro Pro Pro Pro Pro Pro Leu Pro Pro Pro Pro Pro Pro Pro Pro
        35                  40                  45

Leu Pro Pro Leu Pro Leu Pro Pro Leu Lys Lys Arg Gly Asn His Ser
    50                  55                  60
```

```
Thr Gly Leu Cys Leu Leu Val Met Phe Phe Met Val Leu Val Ala Leu
65                  70                  75                  80

Val Gly Leu Gly Leu Gly Met Phe Gln Leu Phe His Leu Gln Lys Glu
                85                  90                  95

Leu Ala Glu Leu Arg Glu Ser Thr Ser Gln Met His Thr Ala Ser Ser
            100                 105                 110

Leu Glu Lys Gln Ile Gly His Pro Ser Pro Pro Pro Glu Lys Lys Glu
        115                 120                 125

Leu Arg Lys Val Ala His Leu Thr Gly Lys Ser Asn Ser Arg Ser Met
    130                 135                 140

Pro Leu Glu Trp Glu Asp Thr Tyr Gly Ile Val Leu Leu Ser Gly Val
145                 150                 155                 160

Lys Tyr Lys Lys Gly Gly Leu Val Ile Asn Glu Thr Gly Leu Tyr Phe
                165                 170                 175

Val Tyr Ser Lys Val Tyr Phe Arg Gly Gln Ser Cys Asn Asn Leu Pro
            180                 185                 190

Leu Ser His Lys Val Tyr Met Arg Asn Ser Lys Tyr Pro Gln Asp Leu
        195                 200                 205

Val Met Met Glu Gly Lys Met Met Ser Tyr Cys Thr Thr Gly Gln Met
    210                 215                 220

Trp Ala Arg Ser Ser Tyr Leu Gly Ala Val Phe Asn Leu Thr Ser Ala
225                 230                 235                 240

Asp His Leu Tyr Val Asn Val Ser Glu Leu Ser Leu Val Asn Phe Glu
                245                 250                 255

Glu Ser Gln Thr Phe Phe Gly Leu Tyr Lys Leu
            260                 265
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 279 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Gln Gln Pro Met Asn Tyr Pro Cys Pro Gln Ile Phe Trp Val Asp
1               5                   10                  15

Ser Ser Ala Thr Ser Ser Trp Ala Pro Pro Gly Ser Val Phe Pro Cys
            20                  25                  30

Pro Ser Cys Gly Pro Arg Gly Pro Asp Gln Arg Arg Pro Pro Pro Pro
        35                  40                  45

Pro Pro Pro Val Ser Pro Leu Pro Pro Pro Ser Gln Pro Leu Pro Leu
    50                  55                  60

Pro Pro Leu Thr Pro Leu Lys Lys Lys Asp His Asn Thr Asn Leu Trp
65                  70                  75                  80
```

```
Leu Pro Val Val Phe Phe Met Val Leu Val Ala Leu Val Gly Met Gly
                85                  90                  95

Leu Gly Met Tyr Gln Leu Phe His Leu Gln Lys Glu Leu Ala Glu Leu
            100                 105                 110

Arg Glu Phe Thr Asn Gln Ser Leu Lys Val Ser Ser Phe Glu Lys Gln
        115                 120                 125

Ile Ala Asn Pro Ser Thr Pro Ser Glu Lys Lys Glu Pro Arg Ser Val
    130                 135                 140

Ala His Leu Thr Gly Asn Pro His Ser Arg Ser Ile Pro Leu Glu Trp
145                 150                 155                 160

Glu Asp Thr Tyr Gly Thr Ala Leu Ile Ser Gly Val Lys Tyr Lys Lys
                165                 170                 175

Gly Gly Leu Val Ile His Glu Thr Gly Leu Tyr Phe Val Tyr Ser Lys
            180                 185                 190

Val Tyr Phe Arg Gly Gln Ser Cys Asn Asn Gln Pro Leu Asn His Lys
        195                 200                 205

Val Tyr Met Arg Asn Ser Lys Tyr Pro Glu Asp Leu Val Leu Met Glu
    210                 215                 220

Glu Lys Arg Leu Asn Tyr Cys Thr Thr Gly Gln Ile Trp Ala His Ser
225                 230                 235                 240

Ser Tyr Leu Gly Ala Val Phe Asn Leu Thr Ser Ala Asp His Leu Tyr
                245                 250                 255

Val Asn Ile Ser Gln Leu Ser Leu Ile Asn Phe Glu Glu Ser Lys Thr
            260                 265                 270

Phe Phe Gly Leu Tyr Lys Leu
        275
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 233 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
Met Ser Thr Glu Ser Met Ile Arg Asp Val Glu Leu Ala Glu Glu Ala
 1               5                  10                  15

Leu Pro Lys Lys Thr Gly Gly Pro Gln Gly Ser Arg Arg Cys Leu Phe
            20                  25                  30

Leu Ser Leu Phe Ser Phe Leu Ile Val Ala Gly Ala Thr Thr Leu Phe
        35                  40                  45

Cys Leu Leu His Phe Gly Val Ile Gly Pro Gln Arg Glu Glu Ser Pro
    50                  55                  60

Arg Asp Leu Ser Leu Ile Ser Pro Leu Ala Gln Ala Val Arg Ser Ser
65                  70                  75                  80
```

```
Ser Arg Thr Pro Ser Asp Lys Pro Val Ala His Val Val Ala Asn Pro
                  85                  90                  95

Gln Ala Glu Gly Gln Leu Gln Trp Leu Asn Arg Arg Ala Asn Ala Leu
             100                 105                 110

Leu Ala Asn Gly Val Glu Leu Arg Asp Asn Gln Leu Val Val Pro Ser
         115                 120                 125

Glu Gly Leu Tyr Leu Ile Tyr Ser Gln Val Leu Phe Lys Gly Gln Gly
     130                 135                 140

Cys Pro Ser Thr His Val Leu Leu Thr His Thr Ile Ser Arg Ile Ala
145                 150                 155                 160

Val Ser Tyr Gln Thr Lys Val Asn Leu Leu Ser Ala Ile Lys Ser Pro
                 165                 170                 175

Cys Gln Arg Glu Thr Pro Glu Gly Ala Glu Ala Lys Pro Trp Tyr Glu
             180                 185                 190

Pro Ile Tyr Leu Gly Cys Val Phe Gln Leu Glu Lys Gly Asp Arg Leu
         195                 200                 205

Ser Ala Glu Ile Asn Arg Pro Asp Tyr Leu Asp Phe Ala Glu Ser Gly
     210                 215                 220

Gln Val Tyr Phe Gly Ile Ile Ala Leu
225                 230
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 205 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Met Thr Pro Pro Glu Arg Leu Phe Leu Pro Arg Val Cys Gly Thr Thr
  1                 5                  10                  15

Leu His Leu Leu Leu Leu Gly Leu Leu Leu Val Leu Leu Pro Gly Ala
             20                  25                  30

Gln Gly Leu Pro Gly Val Gly Leu Thr Pro Ser Ala Ala Gln Thr Ala
         35                  40                  45

Arg Gln His Pro Lys Met His Leu Ala His Ser Thr Leu Lys Pro Ala
     50                  55                  60

Ala His Leu Ile Gly Asp Pro Ser Lys Gln Asn Ser Leu Leu Trp Arg
 65                  70                  75                  80

Ala Asn Thr Asp Arg Ala Phe Leu Gln Asp Gly Phe Ser Leu Ser Asn
                 85                  90                  95

Asn Ser Leu Leu Val Pro Thr Ser Gly Ile Tyr Phe Val Tyr Ser Gln
             100                 105                 110

Val Val Phe Ser Cys Lys Ala Tyr Ser Pro Lys Ala Pro Ser Ser Pro
         115                 120                 125
```

```
Leu Tyr Leu Ala His Glu Val Cys Leu Phe Ser Ser Gln Tyr Pro Phe
    130                 135                 140

His Val Pro Leu Leu Ser Ser Gln Lys Met Val Tyr Pro Gly Leu Gln
145                 150                 155                 160

Glu Pro Trp Leu His Ser Met Tyr His Gly Ala Ala Phe Gln Leu Thr
                165                 170                 175

Gln Gly Asp Gln Leu Ser Thr His Thr Asp Gly Ile Pro His Leu Val
            180                 185                 190

Leu Ser Pro Ser Thr Val Phe Phe Gly Ala Phe Ala Leu
        195                 200                 205
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GCGGCGGGAT CCATGGCTAT GATGGAGGTC CAG 33

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CGCGCGTCTA GAGCTTAGGC AACTAAAAAG GCC 33

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CCGCGCGGAT CCATCATGGC TATGATGGAG GTCC 34

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CGCGCGTCTA GAGCTTAGCC AACTAAAAAG GCC 33

**WHAT IS CLAIMED IS:**

1. An isolated polynucleotide comprising a member selected from the group consisting of:

(a) a polynucleotide having at least a 70% identity to a polynucleotide encoding a polypeptide comprising an amino acid sequence of SEQ ID NO:2;

(b) a polynucleotide which is complementary to the polynucleotide of (a); and

(c) a polynucleotide comprising at least 15 bases of the polynucleotide of (a).

2. The polynucleotide of Claim 1 wherein the polynucleotide is DNA.

3. The polynucleotide of Claim 1 wherein the polynucleotide is RNA.

4. The polynucleotide of Claim 1 wherein the polynucleotide is genomic DNA.

5. The polynucleotide of Claim 2 which encodes the polypeptide comprising amino acids 39 to 281 of SEQ ID NO:2.

6. An isolated polynucleotide comprising a member selected from the group consisting of:

(a) a polynucleotide which encodes a mature polypeptide having the amino acid sequence expressed by the human cDNA contained in ATCC Deposit No. 97448;

(b) a polynucleotide which is complementary to the polynucleotide of (a); and

(c) a polynucleotide comprising at least 15 bases of the polynucleotide of (a) or (b).

7. The polynucleotide of claim 1 comprising the sequence as set forth in SEQ ID No. 1 from nucleotide 1 to nucleotide 1642.

8. The polynucleotide of claim 1 comprising the sequence as set forth in SEQ ID No. 1 from nucleotide 52 to nucleotide 843.

9. A vector comprising the DNA of Claim 2.

10. A host cell comprising the vector of Claim 9.

11. A process for producing a polypeptide comprising: expressing from the host cell of Claim 10 the polypeptide encoded by said DNA.

12. A process for producing a cell which expresses a polypeptide comprising genetically engineering the cell with the vector of Claim 9.

13. A polypeptide comprising a member selected from the group consisting of:

(a) a polypeptide having an amino acid sequence set forth in SEQ ID NO:2; and

(b) a polypeptide which is at least 70% identical to the polypeptide of (a).

14. The polypeptide of Claim 13 wherein the polypeptide comprises amino acid 1 to amino acid 281 of SEQ ID NO:2.

15. A compound which inhibits activation of the polypeptide of claim 13.

16. A compound which activates the polypeptide of claim 13.

17. A method for the treatment of a patient having need of AIM-1 comprising: administering to the patient a therapeutically effective amount of the polypeptide of claim 13.

18. The method of Claim 17 wherein said therapeutically effective amount of the polypeptide is administered by providing to the patient DNA encoding said polypeptide and expressing said polypeptide *in vivo*.

19. A method for the treatment of a patient having need to inhibit AIM-I polypeptide comprising: administering to the patient a therapeutically effective amount of the compound of Claim 15.

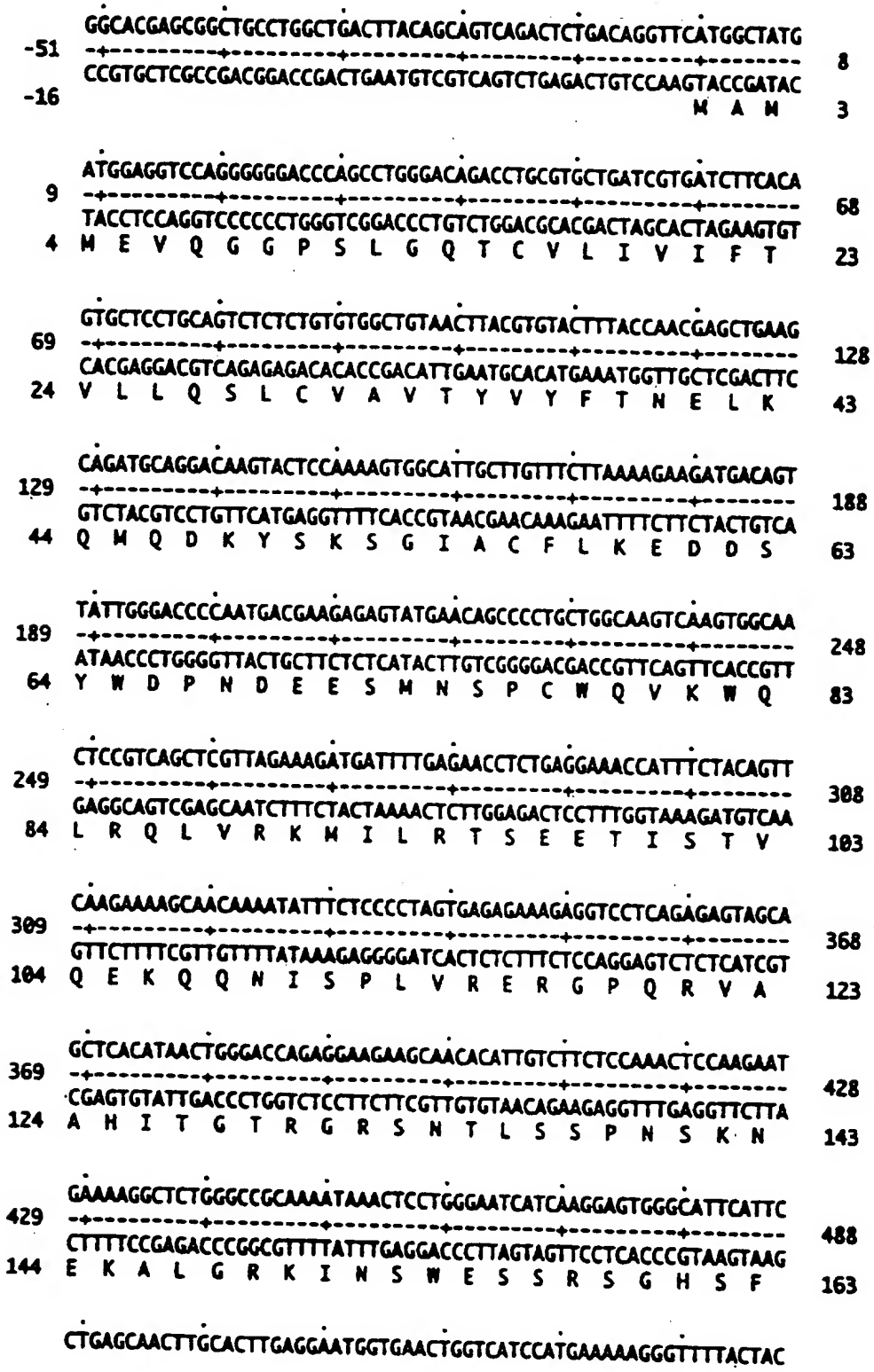
20. A process for diagnosing a disease or a susceptibility to a disease related to an under-expression of the polypeptide of claim 13 comprising determining a mutation in a nucleic acid sequence encoding said polypeptide.

21. A diagnostic process comprising analyzing for the presence of the polypeptide of claim 13 in a sample derived from a host.

22. A method for identifying compounds which bind to and inhibit activation of the polypeptide of claim 13 comprising: contacting a cell expressing on the surface thereof a receptor for the polypeptide, said receptor being associated with a second component capable of providing a detectable signal in response to the binding of a compound to said receptor, with an analytically detectable AIM-I polypeptide and a compound under conditions to permit binding to the receptor; and

determining whether the compound binds to and inhibits the receptor by detecting the absence of a signal generated from the interaction of the AIM-I with the receptor.

## Nucleotide and Amino Acid Sequence of AIM-I



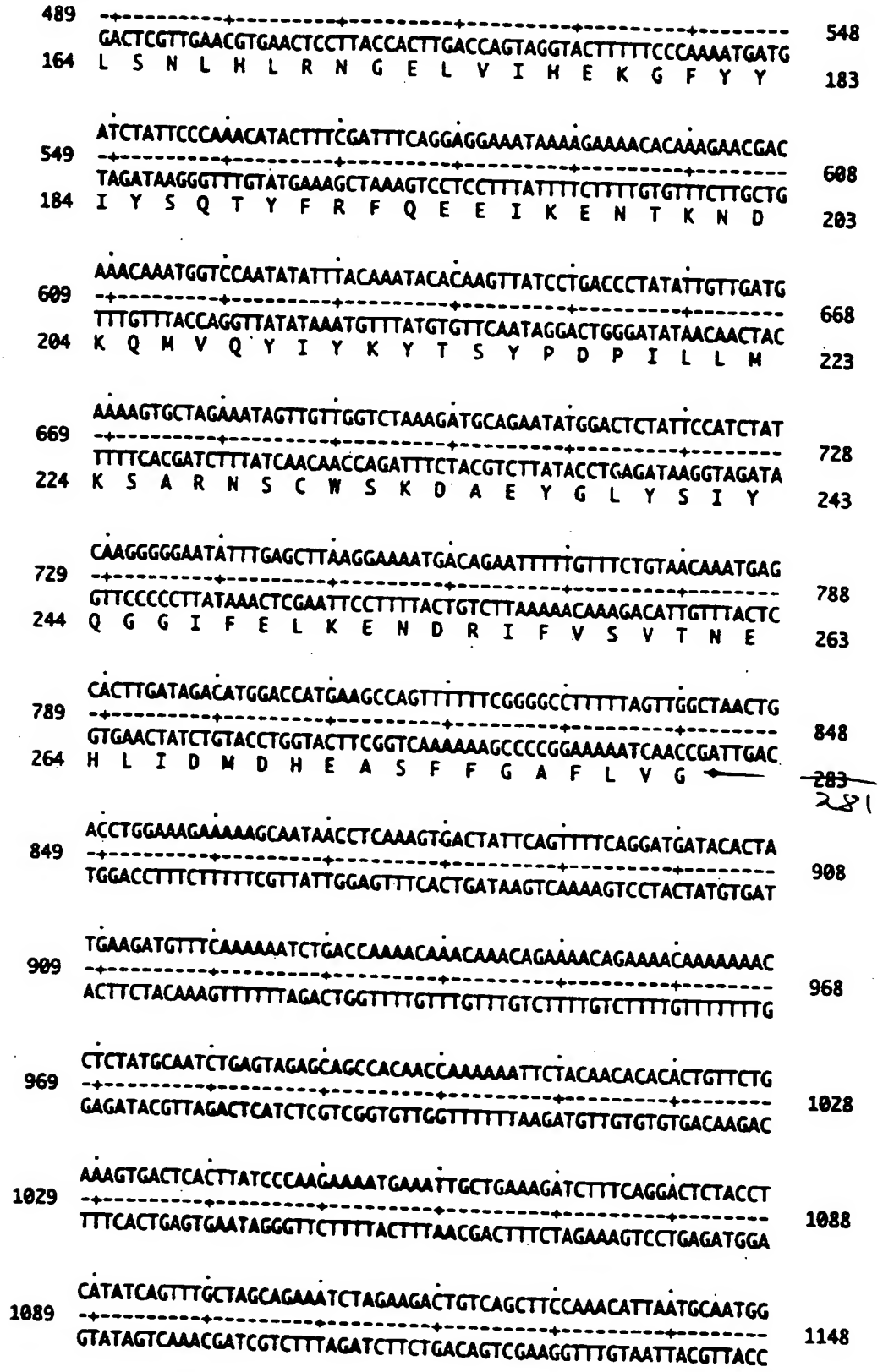

489
GACTCGTTGAACGTGAACTCCTTACCACTTGACCAGTAGGTACTTTTTCCCAAAATGATG
548
164 L S N L H L R N G E L V I H E K G F Y Y 183
ATCTATTCCCAAACATACTTTCGATTTCAGGAGGAAATAAAAGAAAACACAAAGAACGAC
549
TAGATAAGGGTTTGTATGAAAGCTAAAGTCCTCCTTTATTTTCTTTTGTGTTTCTTGCTG
608
184 I Y S Q T Y F R F Q E E I K E N T K N D 203
AAACAAATGGTCCAATATATTTACAAATACACAAGTTATCCTGACCCTATATTGTTGATG
609
TTTGTTTACCAGGTTATATAAATGTTTATGTGTTCAATAGGACTGGGATATAACAACTAC
668
204 K Q M V Q Y I Y K Y T S Y P D P I L L M 223
AAAAGTGCTAGAAATAGTTGTTGGTCTAAAGATGCAGAATATGGACTCTATTCCATCTAT
669
TTTTCACGATCTTTATCAACAACCAGATTTCTACGTCTTATACCTGAGATAAGGTAGATA
728
224 K S A R N S C W S K D A E Y G L Y S I Y 243
CAAGGGGGAATATTTGAGCTTAAGGAAAATGACAGAATTTTTGTTTCTGTAACAAATGAG
729
GTTCCCCCTTATAAACTCGAATTCCTTTTACTGTCTTAAAAACAAAGACATTGTTTACTC
788
244 Q G G I F E L K E N D R I F V S V T N E 263
CACTTGATAGACATGGACCATGAAGCCAGTTTTTTCGGGGCCTTTTTAGTTGGCTAACTG
789
GTGAACTATCTGTACCTGGTACTTCGGTCAAAAAAGCCCCGGAAAAATCAACCGATTGAC
848
264 H L I D M D H E A S F F G A F L V G 283
281
ACCTGGAAAGAAAAAGCAATAACCTCAAAGTGACTATTCAGTTTTCAGGATGATACACTA
849
TGGACCTTTCTTTTTCGTTATTGGAGTTTCACTGATAAGTCAAAAGTCCTACTATGTGAT
908
TGAAGATGTTTCAAAAAATCTGACCAAAACAAACAAACAGAAAACAGAAAACAAAAAAAC
909
ACTTCTACAAAGTTTTTTAGACTGGTTTTGTTTGTTTGTCTTTTGTCTTTTGTTTTTTG
968
CTCTATGCAATCTGAGTAGAGCAGCCACAACCAAAAAATTCTACAACACACACTGTTCTG
969
GAGATACGTTAGACTCATCTCGTCGGTGTTGGTTTTTTAAGATGTTGTGTGTGACAAGAC
1028
AAAGTGACTCACTTATCCCAAGAAAATGAAATTGCTGAAAGATCTTTCAGGACTCTACCT
1029
TTTCACTGAGTGAATAGGGTTCTTTTACTTTAACGACTTTCTAGAAAGTCCTGAGATGGA
1088
CATATCAGTTTGCTAGCAGAAATCTAGAAGACTGTCAGCTTCCAAACATTAATGCAATGG
1089
GTATAGTCAAACGATCGTCTTTAGATCTTCTGACAGTCGAAGGTTTGTAATTACGTTACC
1148

Figure 1 (con'd)

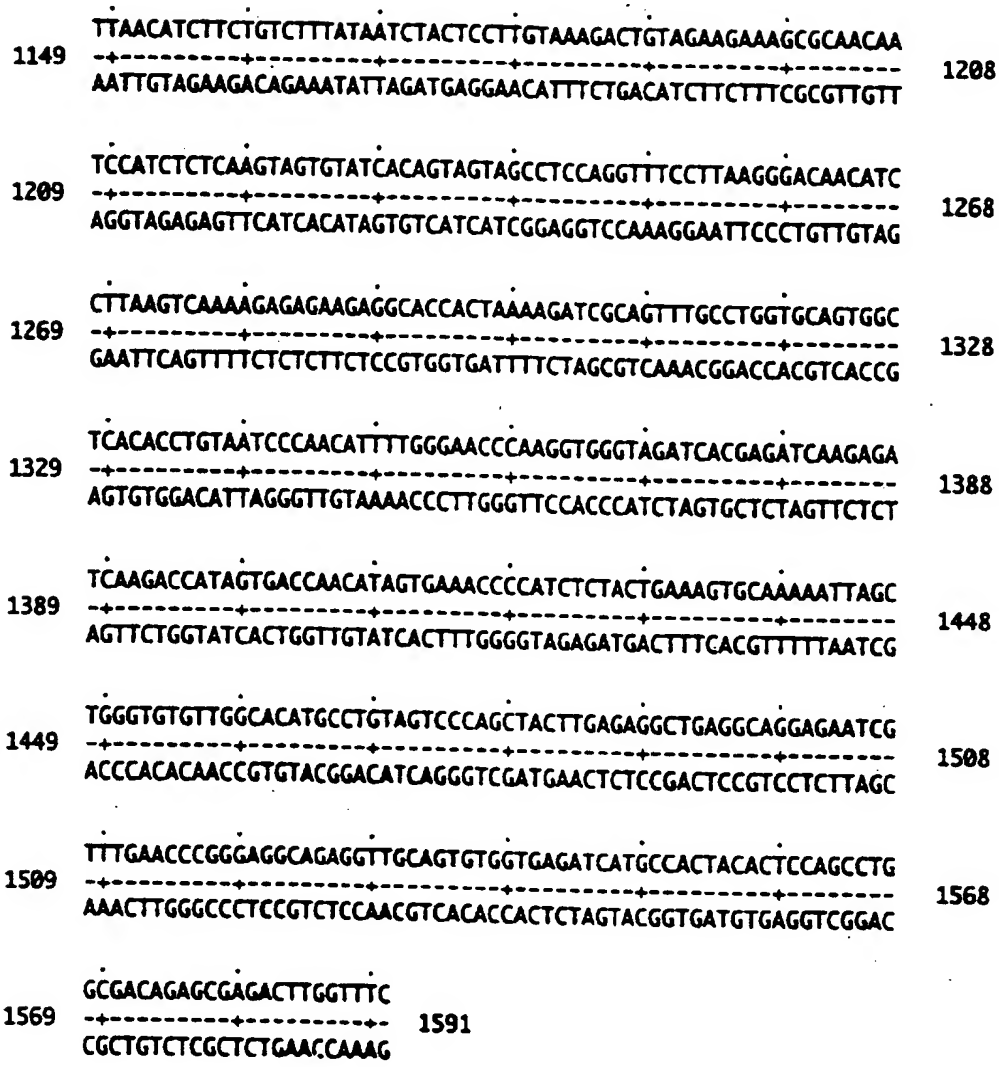
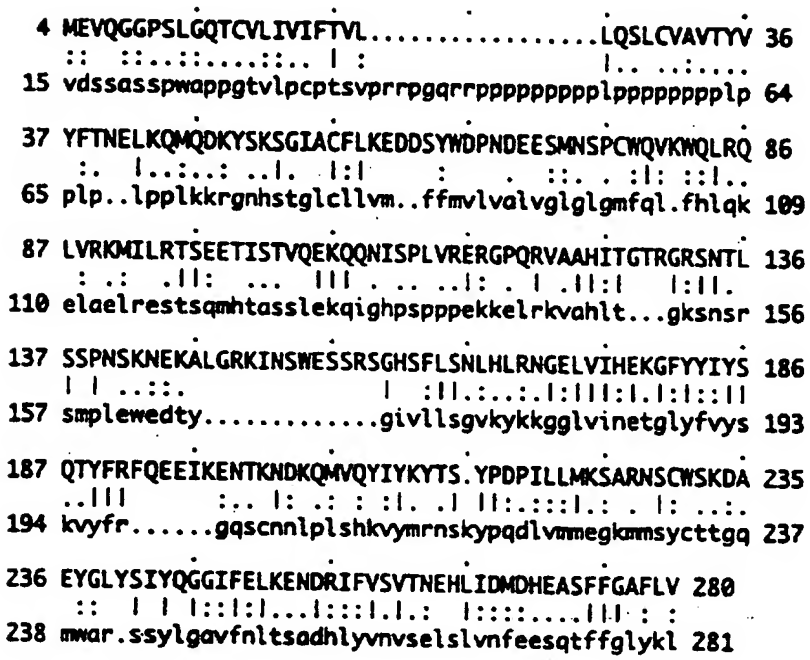


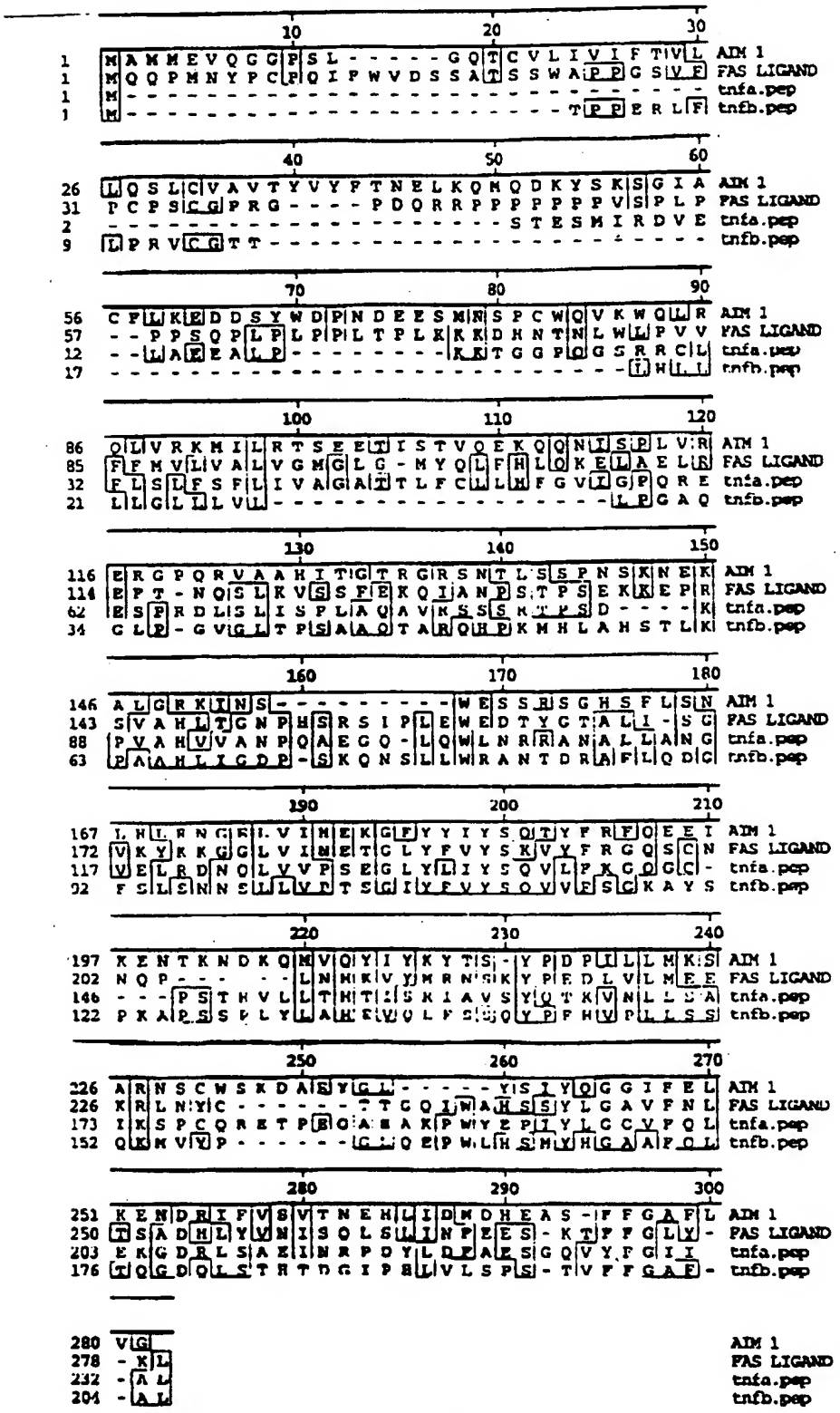
Figure 2

Alignment of AIM-I to Human Fas Ligand
(Similarity = 48.594 % Identity = 22.892 %)



## Figure 3

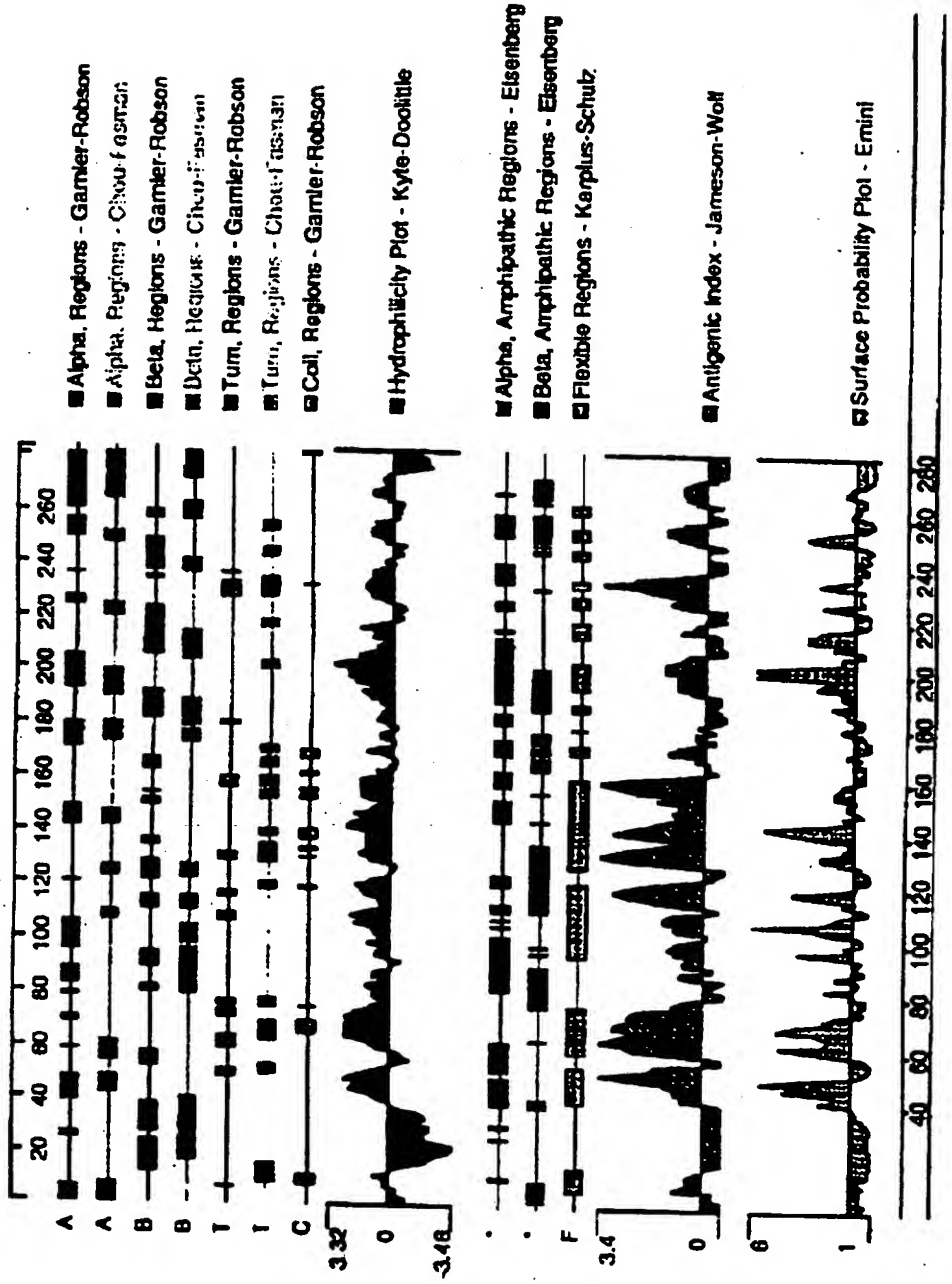
### Alignment Report of AIM-I, hFas Ligand, TNF-α and TNF-β by Clustal Method with PAM250 Residue Weight Table




Decoration 'Decoration #1': Box residues that match the Consensus within 2 distance units.

Figure 4

Regional Analysis of AIM-I Protein



# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/03773

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(6) :C07H 21/00; C07K 14/00
US CL :536/23.1; 530/350

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. 536/23.1, 530/350

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE, EMBASE, BIOSIS, CAPLUS

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| T | PAPOFF et al. An N-Terminal Domain Shared by Fas/Apo-1 (CD95) Soluble Variants Prevents Cell Death in Vitro. The Journal of Immunology. 1996, Vol.156, pages 4622-4630, see entire document. | 1-22 |
| Y | SUDA et al. Molecular Cloning and Expression of the Fas Ligand, a Novel Member of the Tumor Necrosis Factor Family. Cell. 17 December 1993, Vol.75, pages 1169-1178, see entire document. | 1-22 |
| Y | CASCINO et al. Three Functional Soluble Forms of the Human Apoptosis-Inducing Fas Molecule Are Produced by Alternative Splicing. The Journal of Immunology. 1995, Vol.154, pages 2706-2713, see entire document. | 1-22 |

[X] :Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 08 JULY 1996 | 30 JUL 1996 |

| Name and mailing address of the ISA/US | Authorized officer |
|---|---|
| Commissioner of Patents and Trademarks<br>Box PCT<br>Washington, D.C. 20231 | D. CURTIS HOGUE, JR. |
| Facsimile No. (703) 305-3230 | Telephone No. (703) 308-0196 |

Form PCT/ISA/210 (second sheet)(July 1992)*